ETKF−Based Ensemble Perturbations Using Real−time Observations at NCEP

Mozheng Wei[*]

UCAR at NOAA/NWS/NCEP, Camp Springs, MD

Zoltan Toth

NOAA/NWS/NCEP, Camp Springs, MD

Richard Wobus

SAIC at NOAA/NWS/NCEP, Camp Springs, MD

Yuejian Zhu

NOAA/NWS/NCEP, Camp Springs, MD

Craig H. Bishop

Naval Research Laboratory, Monterey, CA

Xuguang Wang

NOAA−CIRES/CDC, Boulder, CO

---

[*] Corresponding author address: Mozheng Wei, NCEP Environmental Modeling Center, 5200 Auth Road, Rm. 207, Camp Springs, MD 20746.

E−mail:Mozheng.Wei@noaa.gov

ABSTRACT

The initial perturbations used for the operational global ensemble prediction system of the National Centers for Environmental Prediction are generated through the breeding method with a regional rescaling mechanism. Limitations of the system include the use of a climatologically fixed estimate of the analysis error variance and the lack of an orthogonalization in the breeding procedure. The ETKF analysis method, as shown by Wang and Bishop, can be used to generate ensemble perturbations that can potentially ameliorate these shortcomings. In the present paper, the ETKF method is tested in an operational environment, with information on the actual distribution and error characteristics of real-time observations. The ETKF method is used with the simplex method (ten individual perturbations are orthogonalized and centered on the control analysis), and compared with the breeding method with five positive-negative pairs of perturbations.

The experimental results show no major difference between the overall performances of the two systems. Though the ETKF initial perturbation variance is shown to respond to changes in the observational network, with only 10 perturbations, the ETKF method cannot represent as well the variations on the global scale in analysis error variance as the breeding method with geographical rescaling. The effective degrees of freedom (EDF) within the subspace of the two sets of perturbations is very similar, with the ETKF method producing only slightly higher number of EDF. As a new result, it is shown that both the ETKF and the bred perturbations exhibit a high degree of continuity in time from one cycle to the next, a favorable property in some practical applications. Additionally, it was found that while the ETKF perturbations explain better the spatio-temporal variations in short range forecast error variance, the error covariance structure of such forecasts are better captured by the bred perturbations. How much benefit the ETKF method can provide when using a larger than 10-member ensemble will be explored in future experiments.

2

1. Introduction

It is well-known that the weather system is chaotic, and its predictability is severely limited by both initial and model-related errors. A feasible way to improve a single, deterministic forecast is to use ensemble forecasting. Ensemble forecasts start from a set of different states that are sampled from a probability density function which is approximated using a finite sample of initial perturbations. However, how to best generate these initial perturbations for an ensemble forecasting system is still an open question.

At the European Center for Medium-Range Weather Forecasts (ECMWF), singular vectors (SVs) are used to identify the directions of fastest forecast error growth for a finite time period (Buizza and Palmer,1995; Molteni *et al.* 1996). Instead of using SVs, the National Centers for Environmental Prediction (NCEP) uses bred vectors (BVs) to sample amplifying analysis errors through breeding cycles that are similar to data assimilation cycles (Toth and Kalnay,1993; 1997). However, both SVs and BVs cannot accurately represent the true uncertainties in analysis as a good ensemble forecast system expects. A comparison of performance between the ECMWF and NCEP ensemble forecast systems was described in Zhu *et al.* (1996), and a more recent comparison can be found in Wei and Toth (2003).

Another major method is the perturbed observation (PO) approach developed at the Meteorological Service of Canada (MSC) (Houtekamer et al. 1996; Houtekamer and Mitchell,1998). The PO approach generates initial conditions by assimilating randomly perturbed observations using different models in a number of independent cycles. The initial perturbations generated by the PO method are more representative of analysis uncertainties in comparison with SVs and BVs. A comprehensive summary of the current methodologies and performance of the three ensemble forecast systems from ECMWF, MSC and NCEP can be found in Buizza *et al.* (2005).

In this paper, we explore a method proposed by Wang and Bishop (2003) (referred to as WB) to generate the initial perturbations for global ensemble forecasts. The method is based on an Ensemble Transform Kalman Filter (ETKF) put forward by Bishop *et al*. (2001). The ETKF was initially applied to the adaptive sampling problem; for example, Majumdar et al. (2001, 2002). Later Wang and Bishop (2003) showed how it could be used to generate ensemble perturbations without having to perform data assimilation while Etherton and Bishop (2004) showed how ETKF ensemble perturbations enabled a highly efficient Hybrid data assimilation scheme. As in Wang and Bishop (2003), in this study, the ETKF is not used to carry out any data assimilation, although ETKF formulation is derived from ensemble Kalman filter theory which is also for data assimilation. Here, the ETKF is only used to transform forecast perturbations into analysis perturbations in a manner consistent with the Kalman Filter error covariance update equation. The ETKF transformation procedure requires as input the locations and error covariances of observations. The ETKF cycling procedure is similar to breeding cycles in that both schemes create analysis perturbations from forecast perturbations. The observational values are used only in computing inflation factors for adjusting the magnitudes of analysis perturbations. The ETKF analysis perturbations are then added to the analysis field produced by the NCEP operational data assimilation system (Parrish and Derber,1992), instead of the one that could be produced by ETKF-based data assimilation. The reason for using the NCEP operational analysis field rather than an analysis based on some sort of ensemble Kalman filter is because the ETKF and other related ensemble-based data assimilation schemes (described below) have not yet been proven superior to the existing NCEP system. The question of whether such ensemble based data assimilation schemes, including ETKF, can generate a good analysis with real observations is being pursued by a few major organizations (see discussion section).

4

WB compared the performance of the ETKF and breeding-based ensemble forecast systems. They showed that the ETKF ensemble produces better results than the breeding method in their experimental setup. However, their experiments were conducted in a simplified environment with an idealized observation system. It would be very interesting to understand how an ETKF-based ensemble forecast system works in an operational environment with real observations. Here are some major differences between WB and our experiments: *(i)*, the two models are different, our NCEP GFS model has a higher resolution (T126L28) than the WB NCAR CCM3 model (T42L18) and we use fewer ensemble members (10) than WB (16). *(ii)*, the observations in our case are real-time, at real locations and the number varies greatly with time, while WB fixed the number of observations and selected the closest model grid points as the observation positions. Thus, the observational operator in WB is simplified. In fact, the accurate computation of the observational operator is one of the major challenges in an operational data assimilation system. (*iii*), our observations can be at any level and irregularly distributed, while WB's are assumed to be at only 3 pre-specified levels. *(iv)*, our observational values are real for calculating inflation factors, while WB used re-analysis data as observations. *(v)*, our observation errors vary spatially and temporally, while WB computed the RMS with re-analysis data as the observational errors. As a matter of fact, WB used only two fixed values for temperature and wind observation error variances, respectively. (vi) In WB's comparison, the magnitude of the initial ensemble spread for the breeding and ETKF have similar values on globally averaged basis, whereas the current comparison does not satisfy this constraint.

Since, in the limit of a very small ensemble (2 members), the ETKF becomes equivalent to the breeding technique with no masking, the question of ensemble size is critical in any comparison of breeding and ETKF ensemble generation techniques. Wang and Bishop's (2003) experiments showed that an 8 member ETKF ensemble was not large enough to resolve geographical fluctuations in

observational density. If limited computational resources limited ones ensemble size to 8 members then one would have had to apply some sort of masking (Toth and Kalnay, 1997) technique to Wang and Bishop's ETKF perturbations to reasonably represent the effect of observational density fluctuations on forecast error variance. Wang and Bishop (2003) did not apply masking to their perturbations because they found that increasing the ensemble size to 16 members was sufficient to crudely resolve the major fluctuations in observational density present in their simulated observational network. One of the objectives of this paper is to investigate whether a relatively small 10 member ETKF ensemble with no masking can outperform a similarly small breeding ensemble with masking. The choice of 10 members is motivated by the simple fact that, currently, NCEP is running a 10 member operational breeding ensemble.

The results from our experiments offer the first test as to how the ensemble based Kalman filter works in an environment that is close to operations, with real-time observations. The comparative evaluations of the ETKF and breeding methods will include the impact of observations in different spaces, such as local, observational, 2-D and 3-D grid point spaces. The amplification factors and effective degrees of freedom of the subspaces spanned by the ETKF and breeding perturbations are compared.

Although the ETKF is not used for data assimilation in this study, the method of generating analysis perturbations (not analysis fields) from forecast perturbations is based on data assimilation principles. In fact, ETKF is one variant of ensemble-based Kalman square-root filters (Tippett *et al*. 2003). Other closely related variants of ensemble-based Kalman filters are the Ensemble Adjustment Kalman Filter (EAKF) and Ensemble Square-root Filter (EnSRF) proposed by Anderson (2001) and Whitaker and Hamill (2002), respectively. A local Ensemble Kalman Filter (LEKF) was proposed by Ott *et al*. (2004) (also see Szunyogh *et al*. 2004). All these methods (ETKF, EAKF, EnSR and LEKF) are deterministic solutions of ensemble Kalman filters, while the PO method is a

6

stochastic solution (Houtekamer and Mitchell,1998; Burgers et al. 1998). Lorenc (2003) has reviewed and compared different ensemble Kalman filters (such as ETKF, EAKF, EnSR, and perturbed observation method) and 4D-VAR for data assimilation.

The paper is organized as follows: Section 2 provides a brief basic description of the ETKF formulations. Also in this section, the experimental setup is described together with the real-time observation data. Section 3 presents the major results from ETKF comparisons with the NCEP operational bred perturbation-based ensemble system. This includes the impact of observations (including a small number of observations such as from Winter Storm Reconnaissance data) on the spread of ensembles, variance distributions, effective degrees of freedom of subspaces spanned by the ensemble perturbations from both systems, amplification factors and optimally combined orthogonal perturbations, and results from comparing ensemble perturbations with forecast error patterns and variances including PECA (Perturbation versus Error Correlation Analysis). Discussion and conclusions are given in Section 4.

2. Methodology

2.1. Basic formulation.

The initial perturbations of the NCEP global ensemble forecast system are generated by a breeding method. This method is well established, widely used and well documented. The operational implementation at NCEP can be found in Toth and Kalnay (1993, 1997). More results and documents are available at the NECP ensemble forecast web site at http://wwwt.emc.ncep.noaa.gov/gmb/ens/index.html.

The ETKF formulation (Bishop *et al*. 2001) is based on the application of a Kalman filter, with the forecast and analysis covariance matrices being represented by k forecast and k analysis perturbations. It is one of the solutions from Kalman filter theory (Anderson 2001; Whitaker and Hamill 2002;

Tippett *et al.* 2003). More details can be found in Bishop *et al.* (2001), Wang and Bishop (2003), Bishop (2003) and Wang *et al.* (2004). Let

$$\mathbf{Z}^f = \frac{1}{\sqrt{k-1}}[\mathbf{z}_1{}^f, \mathbf{z}_2{}^f, \ldots, \mathbf{z}_k{}^f], \qquad \mathbf{Z}^a = \frac{1}{\sqrt{k-1}}[\mathbf{z}_1{}^a, \mathbf{z}_2{}^a, \ldots, \mathbf{z}_k{}^a], \qquad (1)$$

where the *n* dimensional state vectors $\mathbf{z}_i{}^f = \mathbf{x}_i{}^f - \mathbf{x}^f$ and $\mathbf{z}_i{}^a = \mathbf{x}_i{}^a - \mathbf{x}^a$ $(i = 1, 2, \ldots k)$ are *k* ensemble forecast and analysis perturbations, respectively. In our experiments, $\mathbf{x}^f$ is the mean of *k* ensemble forecasts and $\mathbf{x}^a$ is the analysis from the independent NCEP operational data assimilation system. Unless stated otherwise, the lower and upper case bold letters will indicate vectors and matrices respectively. The $n \times n$ forecast and analysis covariance matrices are formed, respectively, as

$$\mathbf{P}^f = \mathbf{Z}^f \mathbf{Z}^{fT} \quad \text{and} \quad \mathbf{P}^a = \mathbf{Z}^a \mathbf{Z}^{aT}, \qquad (2)$$

where T indicates the matrix transpose. For a given set of forecast perturbations $\mathbf{Z}^f$ at time *t*, the analysis perturbations $\mathbf{Z}^a$ can be determined by solving the Kalman filter equation

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H} \mathbf{P}^f, \qquad (3)$$

where $\mathbf{R}$ is the $p \times p$ observational error covariance matrix for *p* observational values that are used in NCEP operational data assimilation system, and $\mathbf{H}$ is the observational operator mapping the forecast grid point values onto the observational points. Let's define $\mathbf{A}^f = \mathbf{R}^{-1/2} \mathbf{H} \mathbf{Z}^f$ as a $p \times k$ matrix, so a singular value decomposition of $\mathbf{A}^f$ can be given by

$$\mathbf{A}^f = \mathbf{U} \mathbf{\Gamma}^{1/2} \mathbf{C}^T. \qquad (4)$$

The ETKF solution is $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T}$, where $\mathbf{T} = \mathbf{C}(\mathbf{\Gamma} + \mathbf{I})^{-1/2}$, $\mathbf{C}$ contains column orthonormal right singular vectors ($\mathbf{c}_i$) and $\mathbf{\Gamma}$ is a diagonal matrix containing squared singular values (λ_i) of $\mathbf{A}^f$; that is, $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k]$ and

8

$\mathbf{\Gamma} = diag(\lambda_1, \lambda_2, \ldots\ldots, \lambda_k)$. Although the forecast perturbations are, by definition, centered about the ensemble mean, i.e. $\sum_{i=1}^{k} \mathbf{z}_i^{\,f} = 0.0$, the analysis perturbations produced by the ETKF defined above are not centered around the analysis ($\sum_{i=1}^{k} \mathbf{z}_i^{\,a} \neq 0.0$). A simple transformation that will preserve $\mathbf{P}^a$ and center the analysis perturbations about the analysis is the simplex transformation first proposed by Purser (1996) (see also Julier,1998; Julier and Uhlmann,1996; Wang *et al*. 2004). $\mathbf{C}^T$ is one of the solutions of this transformation. A proof was shown in Wang *et al*. (2004) Hence, $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T}\mathbf{C}^T$ will be used as our initial analysis perturbations for the next cycle forecasts.

Since the number of ensemble members is too small compared with the nominal degrees of freedom of model state space and since model error is neglected, the analysis error covariance is greatly underestimated by the covariance of the transformed ensemble. Therefore, it is necessary to inflate the analysis perturbations. The inflation method proposed by Wang and Bishop (2003) assumes that the *global* sum of the squares of the differences between a forecast at time t and observations at time t does not depend whether the forecast was initialized at 0Z, 6Z, 12Z or 18Z. It also assumes that the number, quality and location of observations is similar at 0Z, 6Z, 12Z and 18Z. While none of these assumptions are met in an operational system, one of the aims of this paper is to see whether the ETKF can outperform breeding even when the method of defining the inflation factor is ill-posed. Further details of this inflation procedure can be found in Wang *et al*. (2004).

2.2. Experimental setup

Our experiments run from Dec 31, 2002 to Feb. 17, 2003, however, our study will focus on the 32-day period from 01/15/2003 to 2/15/2003. There are 10 ensemble

members in both the ETKF and breeding-based systems. The observations used are from the conventional data set in the NCEP global data assimilation system. This conventional data set contains mostly rawinsonde and various aircraft data, and wind data from satellites. Both the ETKF and breeding ensembles are cycled every 6 hours in accordance with the NCEP data assimilation system, in which new observations are assimilated in consecutive 6-hour time windows centered at 00, 06, 12 and 18 UTC. This is the only difference between our experimental breeding system and the NCEP operational system at the time of the experiments. Please note that the operational breeding system at NCEP was later upgraded to a 6-hour cycle.

The number of observations depends on the observation and telecommunication procedures and generally changes from one cycle to the next. As an example of the data coverage, the distributions of temperature and wind observations at 00Z on January 19, 2003 are shown in Figs. 1 and 2, respectively.

Fig. 1a shows the horizontal locations of temperature data below 500mb, whereas Fig. 1b shows the numbers of temperature observations between different vertical levels. For temperature, the data dense regions are North America, Western Europe and South-East Asia. Shown in Figs. 2a and 2b are the distributions of wind observations that are influenced by satellite passes over different areas. The total numbers of observations over the globe and the Southern Hemisphere only, at different cycles for this time period, are shown in Figs. 3a and 3b. As usual, the number of observations over the Northern Hemisphere is much larger than that over the Southern Hemisphere.

In the following two sections, we will present the results as described in the Introduction.

3. Results from a comparison between ETKF and breeding ensembles

3.1. Impact of observations on the ensemble spread

One of the main attractions of using ETKF ensemble generation is that it allows ensemble variance to reflect the impact of variations in observational density on analysis and forecast error variance provided the ensemble is large enough. To measure the impact of observations on ensemble variance, we will use a total energy measure of ensemble variance. This measure is computed from winds and temperature using

$$E(i,j,k) = \frac{1}{2}[u^2(i,j,k) + v^2(i,j,k) + \frac{C_p}{T_r}T^2(i,j,k)],$$ (8)

where i, j, k are indices for horizontal and vertical directions in the grid point space; u, v, T are the wind components (east-west, north-south) and temperature perturbations, respectively; $C_p = 1004.0 \, J \, kg^{-1} \, K^{-1}$ is the specific heat at constant pressure for dry air and T_r is the reference temperature, following the definition used in Rabier et al. (1996), Wang and Bishop (2003) and Wei and Toth (2003).

Fig. 4 shows global distributions of the energy spread of analysis perturbations and the ratios of the analysis and forecast spread averaged over all levels for both ETKF (left panel) and breeding (right panel) ensembles. For the ETKF ensemble (Fig. 4a), the energy spread of analysis perturbations at each of these levels (only averaged is shown) in the Northern Hemisphere is generally lower than that in the Southern Hemisphere, particularly in the North America and Euro-Asia regions, due to the larger number of observations in these regions (see Figs. 1 and 2). The lowest energy spread is shown in the tropics where the atmosphere is generally more stable.

A clearer picture of the impact from observations is given by the ratio of the analysis and forecast spread. This is shown in Fig. 4c. This ratio represents the rescaling factor from the forecast to analysis spread. In North America, Asia and Europe, where there are more data, the rescaling factors are low. In the Southern Hemisphere, the values of rescaling factors in the areas

11

which are covered by the satellite data are lower than in the areas which are missed by the satellites. The energy spread distributions of analysis perturbations from breeding ensembles, shown in Fig. 4b, do not show the observation impact because the observations are not used. The rescaling factors in breeding are designed empirically from climatology data, with lower scaling factors in the North America and Euro-Asia regions where traditionally there are more observations. More details can be found in Toth and Kalnay (1993, 1997). The rescaling factors in the breeding ensemble are particularly low in North America and Europe. One noticeable difference is that the ETKF rescaling factor distribution is noisier than that in the breeding ensemble. This noise is reminiscent of a similar plot shown in Wang and Bishop (2003) for their 8 member ETKF ensemble but not of the plot corresponding to Wang and Bishop's (2003) 16 member ETKF ensemble. Thus, the noisiness of this plot suggests that with only 10 members the ETKF ensemble might benefit from some sort of masking.

To see the vertical distributions of energy spread, we average the energy spread at all grid points at each level. In Fig. 5a, we show the vertical distributions of energy spread for analysis (solid) and forecast (dotted) perturbations, and recalling factors (dashed) from both ETKF (thick lines) and breeding (thin lines) ensembles. In both ensemble systems, analysis and forecast perturbations have relatively larger energy spreads between 600mb and 200mb. However, the averaged rescaling factors remain very uniform at all levels. The average values of both analysis and forecast perturbation spreads, over all levels, are larger in the ETKF ensemble than in the breeding ensemble. They are 2.172 and 2.222 for ETKF analysis and forecast perturbations, respectively, while for the breeding ensemble these values are 1.602 and 1.694. This may contribute to the fact that bred perturbations grow faster than the ETKF perturbations in most cases. This will be discussed in detail in the next sections.

12

Fig. 5b shows the energy spread distribution of analysis and forecast perturbations along the latitude for both ensemble systems. Unlike the distribution in the vertical direction in Fig. 5a, the latitudinal distributions of energy spread from the two ensemble systems are quite different. Generally, the ETKF ensemble has a lower energy spread in the tropics, while the breeding ensemble has a lower spread mainly in the Southern Hemisphere where the ETKF ensemble has larger spread.

3.2. Impact of WSR data
Having studied the impact from a large number of observations in the above subsection, we will look at any signals from a small number of observations. Winter Storm Reconnaissance (WSR) data for a few days will be used to see if there is any influence from WSR data.

We ran 10 ensemble members at 00Z, 06Z, 12Z and 18Z, the same size as the NCEP operational ensemble system at the time of the experiments. To test the impact of observations, we re-ran the ETKF experiments with slightly different observation data at particular times. In the new experiments, we removed the WSR data at 00Z on Jan. 19, 26, 31 and Feb. 01, 03, 08, and 09, 2003. Details about 2003 WSR data can be found at http://wwwt.emc.ncep.noaa.gov/gmb/targobs/target/wsr2003.html. Each experiment started from the exact same initial conditions as the original experiments at the previous cycle (i.e. 6 hours earlier). The new analysis perturbations on these 7 days, at 00Z without WSR data, will be compared with those with the WSR data. On each day at 00Z, there are about 20 observations. Thus, in each of the 7 cases, the difference between the experiments without and with WSR data will exactly reflect the impact of only 20 observations. The average results of these 7 cases are shown in Fig. 6.

Fig. 6 shows the differences between the two experiments without and with WSR data for the vertically-averaged analysis spread for temperature (Fig. 6a) and wind (Figs. 6b). The differences of the ratios between analysis and forecast

13

spreads from the two experiments are shown in Figs. 6 (c) and (d) for temperature and wind, respectively. The black crosses indicate the locations of WSR data. It is clear that when WSR data are removed, analysis perturbations are larger over the region that the WSR was taken. Indeed, WSR data reduced ensemble analysis variance by 1-2% for these 7 cases with just a 10 member ensemble. These results demonstrate how increasing observational density decreases ETKF ensemble variance. Note that in some areas outside the WSR data region, primarily near equator, there is some noise. Convection near the tropics is more active than in other regions, and any differences, including slightly different initial conditions, which might come from the global model integration scheme will amplify quickly.

3.3. Variance distribution

Wang and Bishop's (2003) results indicated that the ETKF maintains significant variance in a substantially larger number of directions than breeding. Here, we investigate this hypothesis for the case of a small 10 member ensemble and real-time observations.

The forecast and analysis covariance matrices in normalized observational space are $\mathbf{A}^f\mathbf{A}^{f^T}$ and $\mathbf{A}^a\mathbf{A}^{a^T}$, respectively, where $\mathbf{A}^a = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{Z}^a$, and $\mathbf{A}^f$ is defined in Section 2. The variances in different eigen-directions are represented by the corresponding eigenvalues of the covariance matrices. Fig's 7a and 7b show the averaged eigenvalues of $\mathbf{A}^f\mathbf{A}^{f^T}$ (6-hour forecast covariance matrix in normalized observational space) over the 32-day test period for the ETKF ensemble and breeding ensemble, respectively. In both schemes, there are only 9 independent directions out of 10 ensemble members, since the initial perturbations are centered around the analysis.

Fig's 7a and b show that, as in Wang and Bishop (2004), the eigenvalue spectrum of the ETKF ensemble is significantly flatter than that for the

14

breeding ensemble if all 9 non-zero eigenvalues are considered. However, the last four eigenvalues of breeding are close to zero only because, by construction, the breeding ensemble is initialized with 5 pairs of identical but oppositely signed initial perturbations. As such, it is appropriate to note that the 1st and 5th ETKF eigenvalues are respectively 3.26x10^4 and 2.2x10^4, while the 1st and 5th eigenvalues of breeding are respectively 4.4x10^4 and 0.8x10^4. Hence, even when only the first 5 eigenvalues are considered the eigenvalue spectrum of the ETKF is considerably flatter than that of the eigenspectrum from the operational breeding scheme.

A quantitative measure of the flatness of the spectrum is a measure of the degrees of freedom of the subspace spanned by the ensemble perturbations. Here, we use the dimension described in Patil *et al.* (2001). It was called bred dimension by Patil et al. (2001), because the authors studied the subspace spanned by the bred vectors in their paper. A similar definition was used by Bretherton et al. (1999), where it was called the effective number of spatial degrees of freedom. Unlike the matrix rank that counts the number of nonzero singular values, this measure takes account of the relative values of variance in different directions, and removes the ambiguity of small nonzero variances due to, say, computing errors. We believe this definition is useful in measuring the dimensions of subspaces spanned by any vectors. If $\mathbf{z}_i$ $(i = 1, 2, \ldots k)$ are k ensemble perturbations (either forecast or analysis), the matrix $\mathbf{Z}$ can be formed from these perturbations as

$\mathbf{Z} = \dfrac{1}{k-1}[\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k]$. The singular value decomposition of $\mathbf{Z}$ leads to k singular

values ω_i $(i = 1, 2, \ldots k)$. Some of the singular values are zero, and some are non-

zero but small. Then, the dimension can be defined as $d = \dfrac{(\sum\limits_{i=1,k} \omega_i)^2}{\sum\limits_{i=1,k} \omega_i^{2}}$. In this

paper, we call it the *effective degrees of freedom* (EDF) of subspace spanned by the k ensemble perturbations $\mathbf{z}_i \ (i = 1, 2, ...k)$. In general, d is not equal to the rank of $\mathbf{Z}$.

Fig. 7a shows that the EDF of the subspace spanned by the 10 ETKF ensemble forecast perturbations is 8.90, when considering the variation of variances in different directions. It should be noted that the rank of the forecast covariance matrix is 9 when the relative variance values in different directions are not considered. The same time mean variance along different directions in the same normalized observational space for bred vectors is also computed. This is shown in Fig. 7b. As expected, the variances are overwhelmingly on the first 5 bred vectors, and one half of the bred vectors have variances close to zero. Hence, the ETKF spectrum is much more evenly distributed. The EDF in bred vector space is 5.89, which is much lower than that in the ETKF implementation. The main reason for this low dimensionality of bred vector space is that, as mentioned earlier, in the NCEP operational ensemble forecast system the initial bred vectors are in pairs, i.e. a plus/minus strategy was implemented (Toth and Kalnay,1993; 1997). The same strategy was employed in the ECMWF ensemble forecast system, where the singular vectors are added to the analysis in pairs (Molteni et al. 1996). It is expected that the EDF of the subspace spanned by the initial singular vectors is also reduced by half. In the rest of the paper when this plus/minus strategy has clear artificial influence, we will only show the results from 5 members from each system. If we consider only the subspace spanned by the first 5 directions with the largest variances, the EDFs are 4.97 and 4.56 for the ETKF and breeding perturbations, respectively. Thus, under this measure, the difference between the EDFs of the two systems is small. Indeed, in the next subsection (d), it is shown that one can find localized regions in grid point space where the dimension implied by the leading 5 singular values of the bred

16

vector subspace actually exceeds the dimension implied by the leading five singular values of the ETKF subspace. Thus, under Patil et al's (2001) dimension measurement of the flatness of the first 5 eigenvalues of the two ensemble covariances the difference in the flatness of the two eigenvalue spectrums does not appear to be profound. Nevertheless, the relative steepness of the breeding ensembles eigenvalue spectrum over the first 5 eigenvalues relative to the ETKF eigenvalue spectrum is not inconsistent with Wang and Bishop's (2004) hypothesis that, as ensemble size is increased, the eigenvalue spectrum of the ETKF ensemble will become profoundly flatter than the breeding ensemble.

3.4. The effective degrees of freedom of perturbations

As we have seen from Fig. 7, the effective degrees of freedom of the subspace spanned by the 10 bred perturbations are reduced by half at the initial times, due to the symmetrical positive/negative pair strategy implemented in the NCEP global ensemble forecast system. Therefore, it is not fair to compare the effective dimensions of the subspaces spanned by 10 perturbations from the two systems. In this subsection, we study the EDF of subspaces spanned by 5 perturbations, where one is chosen from each pair.

The EDFs of subspaces spanned by 5 analysis (solid) and 5 forecast (dotted) perturbations in 3-D grid point spaces are shown in Fig. 8a, for both ETKF (thick) and breeding (thin), for the same period of time. The values of EDFs for 5 analysis and forecast perturbations spanned in ETKF are 4.719 and 4.697, respectively. For 5 independent breeding perturbations (one from each pair), they are 4.596 and 4.609. It is clear that the EDF from 5 ETKF analysis perturbations is only slightly higher than that from 5 independent breeding perturbations. The EDF of 6-hour forecast perturbations is virtually identical for the ETKF(4.697) and breeding (4.609).

Since EDF estimates are related to the distribution of variance along various directions, it is not surprising that EDF estimates using 3-D grid data are similar to those in observational space. For ETKF, the average EDF of 5 analysis perturbations is 4.719 in 3-D grid point space, while in observational space the EDF is 5.0 (Fig. 7a). The average EDF of 5 forecast perturbations is 4.6797 in 3-D space, and 4.97 in observation space. Thus, the EDFs of analysis and forecast perturbations are reduced by 5.6% and 5.5%, respectively, from observation space to 3-D space for ETKF. This implies that the impact of observations on perturbations in global 3-D grid point space is significant.

For breeding-based ensembles, the EDFs of analysis and forecast perturbations in global 3-D space are also comparable to those in observation space (the observations are not used at all in this experiment). The average EDFs for 5 analysis and forecast perturbations are 4.596 and 4.609, respectively. In ETKF, the average EDF over 32 days for 6-hour forecast perturbations is slightly smaller that for the analysis (Fig. 8a). The situation in the breeding experiments is different, with the EDF of subspace spanned by forecast perturbations being larger than that for the analysis perturbations. After 6-hour forecasts, this kind of extreme structure at the initial times, when the perturbations are in exactly plus/minus pairs, is destroyed.

We also looked at the dimension of subspace spanned by temperature perturbations in horizontal 2-dimensional grid point space at each pressure level. The EDFs of analysis and forecast perturbation subspaces at different pressure levels for the ETKF experiment are displayed in Fig. 8b. As in the perturbations in 3-D space, the EDF of 6-hour forecast perturbations is slightly smaller than that of analysis perturbations. The highest EDF values are found at levels from 500mb to 200mb. This is due to the fact that there are more temperature observations around these pressure levels. For the breeding ensembles (Fig. 8b), the difference between EDFs of the analysis and forecast perturbations at all levels is similar to that in 3-D grid point space. The EDF

18

values from ETKF ensemble are much larger than the breeding's at lower levels. The difference between EDFs of ETKF and breeding ensembles for either forecast or analysis perturbations increases from about 400 mb toward 1000mb level.

In the following, we will look at the EDF distributions at grid points for different pressure levels. Using the method described by Patil et al. (2001), we calculate the EDF of subspaces spanned by 5 analysis perturbations from each ensemble that cover only $(2L+1)(2L+1)$ horizontal grid points, where L is the number of grid points near the central points in each direction. The EDF value from this local subspace is defined as the EDF of the central grid point. The EDF distribution at each level can be calculated by moving the central grid point.

We note that the local EDF depends not only on the number of grid points we choose, but also the number of perturbations we use. In our experiments, the numbers of perturbations are the same for the two systems. To see the dependency of local EDF on the number of grid points, we carry out the experiments for $L = 3, 6, 9, 12, 15$. The results are shown in Fig. 9. Fig. 9a shows the averaged EDF values over all 3-dimensional grid point space for the five experimental cases for both ETKF (square) and breeding (diamond) ensembles. For perturbations in smaller areas, the ETKF ensemble perturbations have higher degrees of freedom than the breeding perturbations, however, the bred perturbations have the advantage in larger sized areas. The EDF values for both systems are nearly the same when $L = 9$.

If we average the EDF values over horizontal directions only at each level, then the vertical distribution of EDF is obtained. This is shown in Fig. 9b, with thick lines for ETKF and thin lines for breeding ensembles. Different line styles indicate different cases with different sizes of the small areas that are covered by the ensemble perturbations. The local EDFs for $L = 3, 6, 9, 12, 15$ are displayed in solid, dotted, dashed, dash-dotted and dash-dot-dotted lines,

19

respectively. In general, the local EDF value varies with the geopotential height for both ensemble systems. One obvious difference between the two systems is that the EDF of the breeding ensemble for last 3 cases with larger local areas (thin dashed, dash-dotted and dash-dot-doted lines) varies much less than its counterpart in the ETKF ensemble.

3.5. Amplification of perturbations

Wang and Bishop's (2003) results indicated that the growth rate of the most rapidly growing linear combination of ETKF perturbations significantly exceeded that of the corresponding optimal combination of breeding perturbations. Later experiments by Wang showed that the growth rates of perturbations in their global model were highly sensitive to the initial amplitude of the perturbations. In particular, they found that perturbation growth rate increased as the size of initial perturbations was diminished. While in Wang and Bishop (2003), the breeding technique was constructed so as to ensure that the breeding perturbations had about the same global amplitude as the breeding perturbations, in the experiments reported here, the ETKF perturbations have significantly larger amplitude than the breeding perturbations (see Fig. 5). Despite this discrepancy, it is of interest to compare the growth rates of the two sets of ensemble perturbations. In addition, Wang and Bishop (2003) never compared the growth rates of individual perturbations from the two systems, but it is of considerable interest to measure this growth. The maximum amplification factor from a linear combination of perturbations is calculated using a method similar to Wang and Bishop (2003) and Bishop and Toth (1999).

Fig. 10 shows the amplification factors (AFs) for different forecast lead times averaged from 00Z Jan 15 to 00Z Feb 15, 2003. The AFs are computed for both the individual perturbations and optimally combined orthogonal perturbations from both ETKF and breeding-based systems. Fig. 10a shows the averaged AFs of 500 mb geopotential height perturbations with thick (ETKF) and

20

thin (breeding) lines. The thick solid line is the averaged AF of all 10 perturbations from the ETKF ensemble, while the thin solid line is the averaged AF for all 10 perturbations in the breeding ensemble. It is clear that the average AF from individual perturbations in the breeding ensemble is larger than that of the ETKF ensemble for short and longer lead times. We only show results out to 2 days, since the calculation of AFs for optimally combined orthogonal perturbations assumes the perturbations are linear.

Shown by dotted lines in Fig. 10a are the maximum AFs of the optimally combined orthogonal perturbations from 5 original perturbations of both systems as a function of lead time. The largest AF of the breeding ensemble is still larger than that of the ETKF ensemble for all forecast lead times shown here. To see the growth rate of each individual perturbation from the two systems, we show the AF for each perturbation at 6- (solid) and 48- (dotted) hour lead times in Fig. 10b. At these two lead times, each breeding perturbation has a larger AF than the corresponding ETKF perturbation.

A likely reason for the bred perturbations to have a larger amplification factors than the ETKF perturbations for 500 mb geopotential height is that, as mentioned previously, the amplification factor is related to the initial perturbation size. The ETKF perturbations have much larger spread at 500 mb height (see Fig. 5a) than bred perturbations. This is also one of the reasons that ETKF perturbations have lower AF values, as shown in Fig. 10. To demonstrate this, we compute the AFs of perturbations from both systems for different regions. Fig. 5b shows that the initial spread of ETKF perturbations are much larger than bred perturbations in the global and Northern and Southern Hemisphere regions, but much smaller in the tropics. We then compare the AF values of perturbations from the two systems for 6-hour lead times in these different regions. Table 1 lists average AF values of all individual perturbations and largest AF from the 5 optimally combined orthogonal perturbations for both ensemble systems in all these regions. In the global,

21

Northern and Southern Hemisphere regions where the ETKF ensemble has a larger spread, the AFs of bred perturbations are larger. However, in the tropics where the ETKF has a smaller initial spread, the AFs of ETKF perturbations are larger.

3.6. Representing forecast error covariance

One measure of the performance of initial perturbations in ensemble forecasting is a direct comparison of the ensemble perturbations with the forecast errors. In this subsection, we use PECA (Perturbation versus Error Correlation Analysis) to study the correlation between ensemble perturbations and forecast errors, as described in Wei and Toth (2003).

The PECA values for the two ensemble systems (solid for ETKF, dotted for breeding) for the global, Northern Hemisphere, Southern Hemisphere and tropical regions are displayed in Figs. 11a, b, c, and d, respectively. As we discussed before, only 5 perturbations are used for a fair comparison. The results for all these regions show that for short forecast lead times (6 to 36 hours), bred perturbations have higher PECA values than the corresponding ETKF perturbations. If we consider data only from every 5th day as independent (not shown), the PECA values for the breeding method for the global and Northern Hemisphere domains for the 6 and 12-hour lead time ranges are higher than for the ETKF method at the 10% (or higher) statistical significance level. The breeding and ETKF systems show similar PECA values beyond 24 hour forecast lead time.

While PECA values indicate the correlations between ensemble perturbations and forecast errors, it is also interesting and important to compare the ensemble variance with the forecast variance. To analyze how well the ensemble variance can explain the forecast error variance, we follow the method used in Majumdar et.al (2001, 2002) and Wang and Bishop (2003). First, we compute the ensemble variance and squared error of temperature at each grid point at the 500 mb pressure level for a 6-hour forecast lead time. A scatter-plot which is not shown can then be drawn by using ensemble (abscissa) and squared forecast errors

22

for all the grid points. We next divide the points into 320 equally populated bins in order of increasing ensemble variance. The ensemble and forecast variances are then averaged within each bin. It is the averaged values of each bin that are plotted (solid) in Fig. 12. Based on these binned values, we carry out a least-square regression and obtain a linear regression line, such as $y = a + bx$. This linear regression line is displayed by a dashed line. The values of b are displayed in the figure. If the number of bins is reduced, it is expected that the curve will be smoother. Shown by a dotted line is the result from 20 bins. The variance relationship between ensemble and forecast is studied for global (top panel), Northern (middle panel) and Southern (bottom panel) Hemispheres. The results for ETKF and breeding ensembles are shown in the left and right panels, respectively.

For a good ensemble system, the forecast error variances are distributed around the ensemble variance. The slope of the linear regression line is close to 45 degrees, i.e. the value of b is close to 1.0. Note, however, that the slope of this line will be highly sensitive to the rescaling/inflation factor used to generate the perturbations. Other things being equal, the slope of the line can be controlled by changing the amplitude of initial perturbations. For the 3 regions (Global, Northern hemisphere, Southern hemisphere) shown here, the slopes of the linear regression lines, based on the larger number of binned values, are relatively larger in the breeding ensemble than in the ETKF ensemble. This seems related to the fact that the breeding ensemble generally has smaller ensemble perturbation magnitudes. The results from 20-bin (dotted line) show that the range of forecast error variance (maximum minus minimum value) explained by the ensemble variance is larger for ETKF (5.03) than breeding (2.77) in the global region (Fig.s 12 (a) and (b)). This shows that the ETKF is better than breeding at being able to distinguish times and locations where forecast errors are likely to be large from the times and locations where

23

forecast errors are likely to be small. For the other two regions, the ranges of forecast variances from ETKF are also slightly larger compared with the breeding ensemble.

3.7. Correlation between analysis and forecast perturbations

In this subsection, we compare each forecast perturbation with the corresponding transformed analysis perturbation. In breeding ensemble, the analysis perturbations are scaled from the 6-hour forecast perturbations. That is, $\mathbf{z}^a_i(j) = \alpha_i(j)\mathbf{z}^f_i(j)$, where $\alpha_i(j)$ is the rescaling factor derived from a mask field for ensemble member i and grid point j in horizontal space. In the case of a single global rescaling factor ($\alpha_i(j) = cons\tan t$) at every cycle, the correlation between analysis and forecast perturbations will be 1.0. In this case, the spatial and temporal variations of analysis errors are not accounted for. The mask used in NCEP ensemble forecast system has been constructed from seasonal average of analysis error variances for each month. On each day, the mask is obtained through linear interpolation of the monthly values. It is expected that the correlation values between $\mathbf{z}^a_i$ and $\mathbf{z}^f_i$ in breeding ensemble will be high.

In ETKF theory, the 6-hour forecast perturbations are transformed into analysis perturbations based on Kalman filter theory taking the observation information into account, such as

$$\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} \mathbf{C}^T = \mathbf{Z}^f \mathbf{C} (\mathbf{\Gamma} + \mathbf{I})^{-1/2} \mathbf{C}^T \qquad (9)$$

The transformation from forecast to analysis perturbations can be described as three steps: First, the forecast perturbations $\mathbf{Z}^f$ are rotated by $\mathbf{C}$, they are then scaled by $(\mathbf{\Gamma} + \mathbf{I})^{-1/2}$. Finally they are rotated again by $\mathbf{C}^T$ which is a simplex transformation. The main purpose of the simplex transformation is to center the transformed perturbations around the analysis field while preserving the analysis covariance. In the first step, the forecast perturbations are

rotated into different directions, while the second step only rescales the rotated perturbations. It can be expected that without the last step simplex transformation $\mathbf{C}^T$, the rotated and scaled perturbations would have low correlation with the original forecast perturbations depending on how much the perturbations are rotated. However with the simplex transformation, the rotated and scaled perturbations are rotated toward the directions that are opposite to the first-step rotation by $\mathbf{C}$. If the eigenvalue distribution $\mathbf{\Gamma}$ is completely flat, the correlation between $\mathbf{Z}^a$ and $\mathbf{Z}^f$ will be 1.0.

Shown in Fig. 13 (a) are the averaged correlation values over 10 members between forecast and analysis perturbations for ETKF (thin) and breeding (thick) ensembles at different times. The correlation between the forecast and analysis perturbations at each level is computed for both ensemble systems. The mean correlation over all levels is indicated in solid line. The results show that the mean correlation in ETKF ensemble over different model levels is consistently higher than that in breeding ensemble although the mean correlation varies with time in both ensemble systems. To see the correlation variation with time at different levels, we show the results for levels at 1000mb, 500mb and 2mb in dotted, dashed and dash-dotted lines respectively. At different pressure levels, the correlation between the corresponding forecast and analysis perturbations changes little for ETKF ensemble. However for breeding ensemble, the correlation at different levels varies more, particularly at the top model level (2mb). This variation with pressure level can be seen more clearly from Fig. 16 (b) which shows the vertical distribution of averaged correlation over time period from Jan 15 to Feb 15 2003. The averaged correlation over the experimental period is almost constant at different pressure levels for ETKF ensemble, while the breeding ensemble shows larger variation at different levels.

The main reason for this extremely high correlation between analysis and forecast perturbations in ETKF is the simplex transformation. Equation (9) shows that the correlation in ETKF ensemble is also influenced by the eigenvalue distribution Γ. The eigenvalue distribution is determined by the number and location of observations and the number of ensemble members. In our experiment, the eigenvalue distribution of forecast covariance matrix in the normalized observational space is shown in diamond in Fig. 7(a). The effective degree of freedom is 8.9 out of 9 independent forecast perturbations. The variances are quite evenly distributed in different directions. The correlation between analysis and forecast perturbations in ETKF ensemble is changed little by this distribution. We notice that WB also showed a similar variance distribution in their model with ideal observations. We can reasonably expect that the analysis perturbations in most ETKF ensembles with simplex transformation have high correlation values with the forecast perturbations although the exact influence from the observations and the number of ensembles is hard to know.

There are some ensemble forecasting systems that require high correlation between analysis and forecast perturbations, such as ensemble systems for forecasting ocean waves and other hydrological systems. Higher correlation indicates more time continuity between the perturbations at different cycles. This nice feature makes ETKF based ensemble system particularly appealing.

4. Discussion and conclusions

In this paper, we have carried out experiments with two ensemble forecast systems based on two different techniques for generating initial perturbations: ETKF and breeding. Results are presented for a 32-day experimental period using the NCEP operational analysis/forecast system, and focusing on the characteristics of analysis and short range forecast perturbations. The purpose of this comparison between the ETKF and breeding ensembles is to see if the

26

ETKF-generated initial perturbations are more responsive to observation distributions and representative of the analysis uncertainties, and whether the performance can be improved.

The detailed properties of ETKF-generated perturbations are thoroughly studied from various aspects, such as the effective degrees of freedom of subspaces spanned by perturbations in local, observational, global 2-D and 3-D grid point spaces, optimally combined orthogonal perturbations with the largest amplification factors. The relative strengths and weaknesses of the two systems are discussed and identified. The results presented in this paper for the first time offer a valuable, comprehensive description of the performance of an ETKF-based ensemble forecast system under a real-time observation environment. The paper can be a useful reference for scientists who intend to use any ensemble-based Kalman filter to develop an ensemble forecast or ensemble-based data assimilation system in an operational environment.

The findings from our experiments are summarized as follows:

- The ETKF method is shown to produce initial perturbations whose variance, as desired, is influenced by variations in data coverage.

This is in contrast to some of current operational techniques such as the breeding technique at NCEP and the SV technique, although other techniques, such as the PO method used at MSC, are expected to produce more similar results. The ETKF is capable of picking up the impact from very small number of observations as shown in Section 3b.

- While the slope of the eigenvalue spectrum of the breeding ensemble covariance matrix is clearly steeper than that of the corresponding ETKF eigenvalue spectrum, the effective degrees of freedom of a subspace spanned by 5 bred perturbations (one from each pair) is very similar to that of a subspace spanned by 5 ETKF perturbations.

- In general, the growth rates of ETKF and bred perturbations are similar.

A good ensemble forecast system requires that the initial perturbations grow reasonably fast enough to match the growth rates of forecast errors. In our experiments, bred perturbations have larger amplification factors over the Northern and Southern extra-tropics than ETKF perturbations, while it is ETKF perturbations that have higher amplification factors in the tropics (see Table 1). This may be explained by the fact that the breeding perturbations are generally smaller than the ETKF perturbations everywhere except the tropics. Both systems are based on the same concept of letting cycled perturbations reflect the dynamics of error growth due to initial errors in analysis cycles.

- The scatter plots of ensemble and forecast variances from all grid points show that the ETKF ensemble distinguishes a larger range of forecast variances.

- PECA calculations indicate that bred perturbations explain a larger portion of short range forecast errors than ETKF for the regions we have tested. For longer lead times from 2 days onward, both systems show similar PECA values.

PECA values quantitatively measure how well the ensemble perturbations match the forecast errors. For longer forecast lead times any perturbations including, ETKF ensemble perturbations, will turn toward the leading Lyapunov vector which is linked to the bred vector.

- Both systems produce temporally consistent perturbation fields.

It is found that ETKF analysis perturbations have a very high correlation with forecast perturbations before the ETKF transformation. The correlation between ETKF analysis and forecast perturbations is even slightly higher than that in a breeding system with regional rescaling. This surprising, but good feature of ETKF perturbations is due to the simplex transformation imposed.

We note that the above findings are from the experiments we have carried out so far. There are still some clear limitations in our study, such as:

28

1. Only 5 perturbations are used to compute the effective degrees of freedom of the subspace spanned by each system in order to have a fair comparison. This is mainly due to the fact that there are only 10 members in NCEP operational ensemble forecast system and perturbations are in positive/negative pairs. This small number of ensemble members may be a disadvantage for the ETKF based ensemble. The performance of ETKF ensemble should be greatly improved if the number of ensemble members is increased.

2. Only the so-called `conventional data' from the NCEP operational data assimilation system have been used. There are a lot more satellite data that have not been used. To include those, more work would be needed.

3. One should note that the number of ensemble members is too small compared with the model state space. Projecting the variances from a large state space onto such a small subspace spanned by the ensemble is a simplification. For a given number of ensembles, it is important to inflate the analysis perturbations properly. The inflation strategy we used needs to be improved for system with real-time observations. At present, how to correctly inflate the analysis variances remains a challenging research issue for the ensemble Kalman filter research community. To avoid the ill-posed inflation factor for the ETKF experiment we conducted, one simple way is to inflate the ETKF initial perturbations so that on globally averaged basis the initial ensemble variance of the ETKF is similar to that of the operational breeding. Alternatively, to handle the problem of varying number of observations, instead of using just the current observations we can also try to use previous two weeks' observations and follow the similar steps in Wang and Bishop (2003) to get the inflation factor. Future work will be done on these experiments.

4. The ETKF analysis uncertainty estimate is not fully consistent with the NCEP operational 3-D VAR analysis uncertainty. The main difference between

the ETKF and NCEP 3D-Var operational data assimilation system is that the background covariance in ETKF is represented by ensemble perturbations, while in 3D-Var it is generated by the NMC method (Parrish and Derber, 1992). It is expected that the background covariance matrix produced by the NMC method is more isotropic than that generated by the ensembles. This is particularly true for our small number of ensemble members. There are two ways of avoiding this. First, you can use ensemble based data assimilation. As described in the Introduction, ETKF, EAKF, EnSR and LEKF all are ensemble-based Kalman filters. A major inter-comparison project has been initiated recently at NCEP in cooperation with the people who derived and formulated these filters at the NOAA Climate Diagnostics Center (NOAA CDC), University of Maryland and National Center for Atmospheric Research (NCAR) (Toth *et al.* 2004). This project is supported by THORPEX (see http://box.mmm.ucar.edu/uswrp). The goal of this project is to compare the performance of each of these ensemble-based data assimilation schemes in an environment with real operational (and more sophisticated) models and data. The results will be compared with the benchmark NCEP operational data assimilation system. Second, you can get the analysis uncertainty information from 3D/4D Var and feed it into the ensemble forecast system. We plan to explore this with respect to breeding techniques in the future. Orthogonalization and simplex transformations can be used to restrain initial perturbation variance. The results will be compared with 40-member ETKF.

REFERENCES

Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* **129**, 2884-2903.

Bishop, C.H., Etherton, B.J. and Majumdar, S. 2001. Adaptive sampling with the ensemble transform Kalman filter. part I: theoretical aspects. *Mon. Wea. Rev.,* **129,** 420-436.

Bishop, C.H. and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci*., **56,** 1748-1765.

Bretherton, C. S., Widmann, M., Dymnikov, V. P.,Wallace, J. M. and Blade, I 1999. The effective number of spatial degrees of freedom of a time-varying filedl. *J. Climate*, **12**, 1990-1999.

Buizza, R. and Palmer, T. N. 1995. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci*. **52**, 1434-1456.

Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, P., Wei, M. and Zhu, Y. 2005. A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems. *Mon. Wea. Rev*. in press.

Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998. On the analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev*., **123**, 1719-1724.

Dee, D. P. 1995. On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev*., **123**, 1128-1196.

Etherton, B.J. and Bishop, C. H. 2004. Resilience of hybrid ensemble/3D-Var analysis schemes to model error and ensemble covariance error. *Mon Wea. Rev.* **130,** 1065-1080.

Houtekamer, P. L., Lefaivrem, L., Derome, J.,Ritchie, H. and Mitchell,H.L. 1996. A sytem simulation approach to ensemble prediction. *Mon. Wea. Rev*., **124,** 1225-1242.

Houtekamer, P.L., and Mitchell, H. L. 1998. Data assimilation using an ensemble
    Kalman filter technique. *Mon. Wea. Rev*., **126**, 796-811.

Julier, S. J. 1996. A general method for approximating nonlinear transformations
    of probability distribution, Elsevier preprint.

Julier, S. J. 1998. Reduced sigma point filters for propagation of means and
    covariances through nonlinear transformations, Elservier preprint.

Lorenc, A. C. 2003. The potential of the ensemble Kalman filter for NWP-a
    comparison with 4D-Var. *J. Roy. Meteor. Soc*., **129**, 3183-3203.

Majumdar, S.J., Bishop,C.H., Szunyogh, I and Toth, Z. 2001. Can an Ensemble
    Transform Kalman Filter predict the reduction in forecast error variance
    produced by targeted observations?. *Quart. J. Roy. Met. Soc.* **127**, 2803-2820.

Majumdar, S. J., Bishop, C.H. and Etherton, B.J. 2002. Adaptive sampling with
    Ensemble Transform Kalman Filter. Part II: Field Program Implementation.
    Mon. Wea. Rev., **130**, 1356-1369.

Molteni, F., Buizza, R. Palmer, T. and Petroliagis, T. 1996. The ECMWF ensemble
    prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc*.,
    122, 73-119.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich,E.J., Corazza,M.,
    Kalnayy, E., Patil, D.J. and Yorke, J.A. 2004. A Local ensemble Kalman filter
    for atmospheric data assimilation. *Tellus*, **56A**, 415-428.

Parrish, D. F. and Derber, J. 1992. The National Meteorological Center's
    spectral statistical-interpolation analysis system. *Mon. Wea. Rev*.,
    120, 1747-1763.

Patil, D. J., Hunt, B. R., Kalnay, E., Yorke, J.A. and Ott, E. 2001. Local low
    dimensionality of atmospheric dynamics. *Phys. Rev. Lett*. **86**, 5878-5881.

Purser, R. J. 1996. Arrangement of ensemble in a simplex to produce given first
    and second-moments, NCEP Internal Report (available from the author at
    Jim.Purser@noaa.gov).

Purser, R. J. and Huang, H-L. 1993. Estimating effective data density in

a Satellite retrieval or an objective analysis. *J. Appl. Meteor*. **32**,
   1092–1107.

Rabier, F., Klinker,E, Courtier, P. and Hollingsworth, A. 1996. Sensitivity of
   forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc*. **122**, 121–
   150

Szunyogh, I., Kalnay,I. and Toth. Z. 1997. A comparison of Lyapunov and optimal
   vectors in a low-resolution GCM, *Tellus*, **48A**, 200–227.

Szunyogh, I., Kostelich, E.J., Gyarmati, G., Hunt,B.R., Ott, E., Zimin, A.V.,
   Kalnay, E., Patil, D.J. and York, J.A. 2004. A local ensemble Kalman filter
   for the NCEP GFS model. AMS Meeting, Seattle, Jan, 11–15.

Tippett, M. K., Anderson, J.L., Bishop,C.H., Hamill, T. and Whitaker, J.S.
   2003. Ensemble square root filters. *Mon. Wea. Rev*.,**131**, 1485–1490.

Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: the generation of
   perturbations. *Bull. Amer. Meteor. Soc*., **174**, 2317–2330.

Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding
   method. *Mon. Wea. Rev*., **125**, 3297–3319.

Toth, Z, Wei, M., Anderson, J.A., Parish, D., Szunyogh,I., Whitaker, J and
   Hamill, T. 2004. Intercomparison of difference ensemble–based data
   assimilation schemes.(NCEP research project)

Wang, X., and Bishop, C.H. 2003. A comparison of breeding and ensemble
   transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci*., **60**,
   1140–1158.

Wang, X., Bishop, C. H. and Julier, S. J. 2004. Which is better, an ensemble of
   positive/negative pairs or a centered spherical simplex ensemble?
   Mon. Wea. Rev. **132**, 1590–1605.

Wei, M. 2000. Quantifying local instability and predictability of chaotic
   dynamical systems by means of local ,metric entropy. *Int. J. of Bifurcation
   and Chaos,* **10**, 135–154.

Wei, M. and Frederiksen, J. S. 2004. Error growth and dynamical vectors during

southern hemisphere blocking. *Nonl. Proc. in Geoph*., **11**, 99–118.

Wei, M. and Toth, Z. 2003. A new measure of ensemble performance:

Perturbations versus Error Correlation Analysis (PECA).

Mon. Wea. Rev., **131**, 1549–1565.

Whitaker, J. S. and Hamill, T.M. 2002. Ensemble data assimilation without

perturbed observations. *Mon. Wea. Rev*., **130**, 1913–1924.

Zhu, Y., lyengar, G., Toth, Z., Tracton, M. S. and Marchok, T. 1996. Objective

evaluation of the NCEP global ensemble forecasting system. Preprints of the

15th AMS Conference on Weather Analysis and Forecasting, 19–23 August 1996,

Norfolk, Virginia, p. J79–J82.

Table 1. Amplification factors of 500mb geopotential height
at 6–hour forecast lead time

Ave amplification factor for all individual perturbations				
	GL	*TR*	*NH*	*SH*
Breeding	*1.112*	*1.282*	*1.105*	*1.103*
ETKF	*1.091*	*1.621*	*1.096*	*1.082*
Max amplification from 5 optimal perturbations				
	GL	*TR*	*NH*	*SH*
Breeding	*1.199*	*1.821*	*1.194*	*1.190*
ETKF	*1.140*	*2.061*	*1.158*	*1.151*

Fig. 1. Temperature observation distributions at 00Z January 19, 2003. (a) the horizontal distribution of data below 500mb; (b) the vertical distribution.

Fig. 2. As in Fig.1, but for wind.

Fig. 3. The number of observations at different cycles during the experimental period (a) over the globe; (b) over the Southern Hemisphere only.

Fig. 4. Vertically averaged global distribution of energy spread of analysis perturbations and the ratios of the analysis and forecast spread, for both ETKF and breeding ensembles with (a) energy spread of ETKF; (b) energy spread of breeding ensemble; (c) ratio of analysis spread/forecast spread for ETKF; and (d) ratio of analysis spread/forecast spread for breeding ensemble.

Fig. 5. Energy spread distributions of ETKF (thick) and breeding (thin) ensemble perturbations (solid: analysis; dotted: forecast). The ratio of analysis/forecast perturbations is indicated by the dashed line. All the values are averaged over the period 01/15 – 02/15, 2003, with (a) vertical distribution as a function of pressure; (b) distribution by latitude.

Fig. 6. The difference of vertically averaged analysis spread for (a) temperature and (b) wind, between two experiments with and without WSR data. The difference in analysis/forecast spread for (c) temperature and (d) wind, between the same two experiments.

Fig. 7. The averaged variance distributions along different eigen-directions of forecast (diamond) and analysis (square) covariance matrices in the normalized observational space, for the: (a) ETKF ensemble; (b) breeding ensemble.

Fig. 8. The effective degrees of freedom of subspace spanned by 5 temperature perturbations (solid: 5 analysis perturbations; dotted: 5 forecast perturbations) from ETKF (thick) and breeding (thin) ensembles for: (a) EDF of subspace spanned by temperature perturbations in 3-dimensional grid point space, at different cycles during the experimental period; (b) EDF of subspace spanned

by 5 temperature perturbations in 2-dimensional grid point space at each pressure level.

Fig. 9. Local EDF for different number of grid points for: (a) averaged local EDF as a function of L, which is described in the text; and (b) local EDF at each pressure level.

Fig. 10. Amplification factors of ensemble perturbations for: (a) the averaged AF from 10 individual perturbations as a function of lead time (solid) and the maximum AF of optimally combined orthogonal perturbations from 5 original perturbations (dotted) (The results from ETKF and breeding ensembles are indicated by thick and thin lines respectively); and (b) the AF of 10 individual perturbations for two forecast lead times (solid: 6-hour, dotted: 48-hour), where the ETKF and breeding ensembles are indicated by triangles and diamonds, respectively.

Fig. 11. The PECA values for ETKF (solid) and breeding (dotted) ensembles from 5 perturbations only. Shown in thick and thin lines are PECA from 5 optimally combined perturbations and the average PECA from 5 individual perturbations.

Fig. 12. Derived ensemble variance and forecast error variances at all grid points for 500mb temperature, for ETKF (left panel) and breeding (right panel); for global (top), Northern (middle) and Southern (bottom) Hemisphere regions. The average value from each of 320 bins is indicated by solid lines. Dotted lines show the results from 20 bins only. The linear regression line from 320 bins is displayed by a dashed line.

Fig. 13. Averaged correlation over 10 members between forecast and analysis perturbations for ETKF (thin) and breeding (thick) ensembles. (a) correlation as a function of time. The mean correlation over all levels, correlations at levels 1000mb, 500mb and 2mb are shown in solid, dotted, dashed and dash-dotted lines respectively. (b) vertical correlation distribution averaged over time.

(a) Temp obs below 500mb (1900/01/03), number: 18328

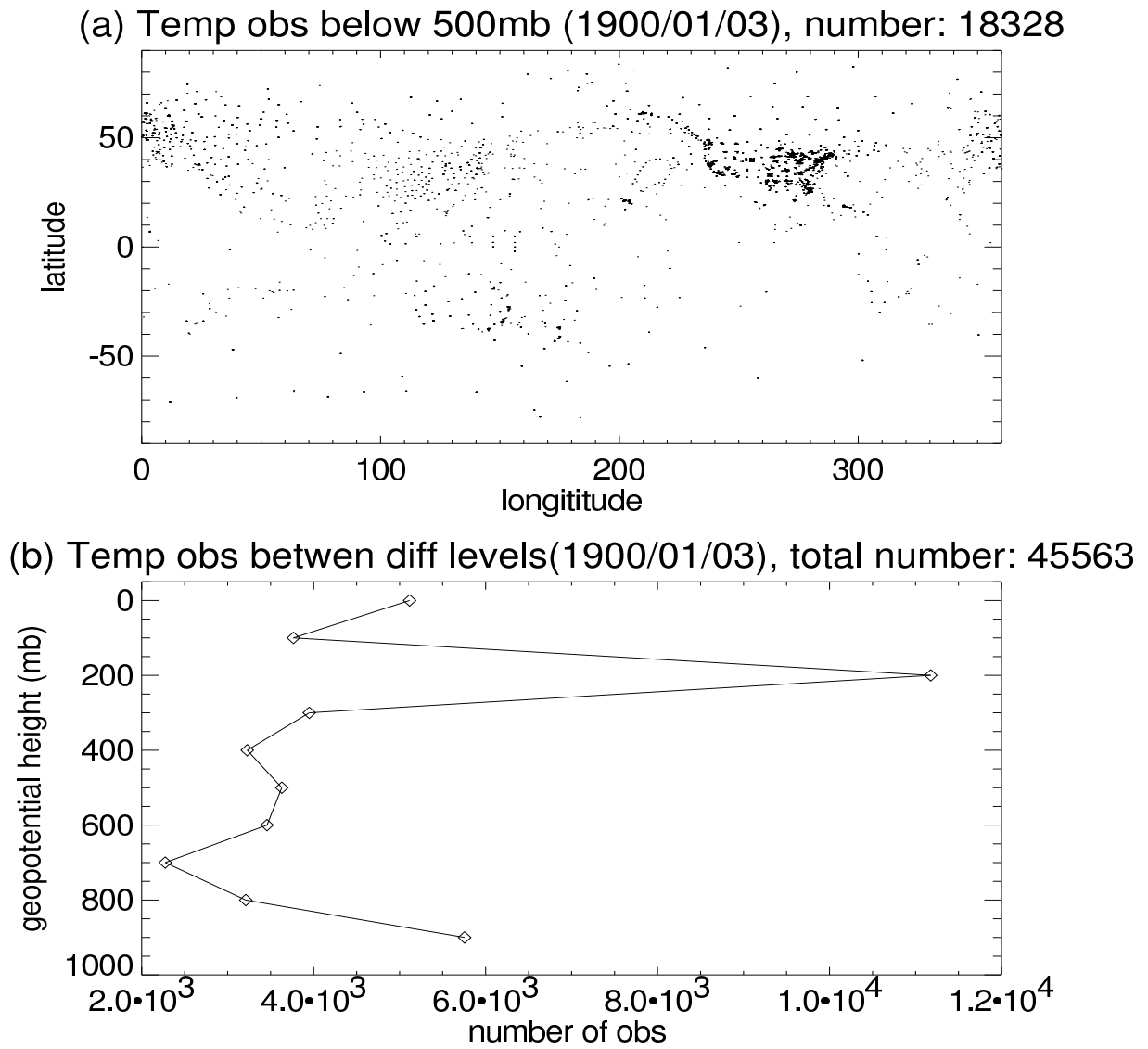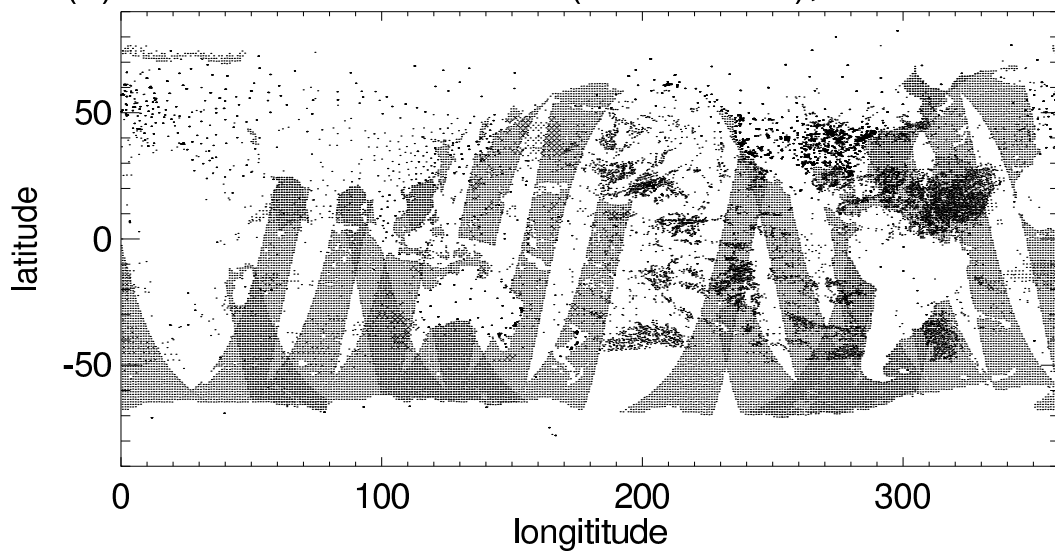## (b) Temp obs betwen diff levels(1900/01/03), total number: 45563

Fig. 1. Temperature observation distributions at 00Z January 19, 2003. (a) the horizontal distributions of data below 500mb; (b) the vertical distribution.

(a) Wind obs below 500mb(1900/01/03), number: 62629

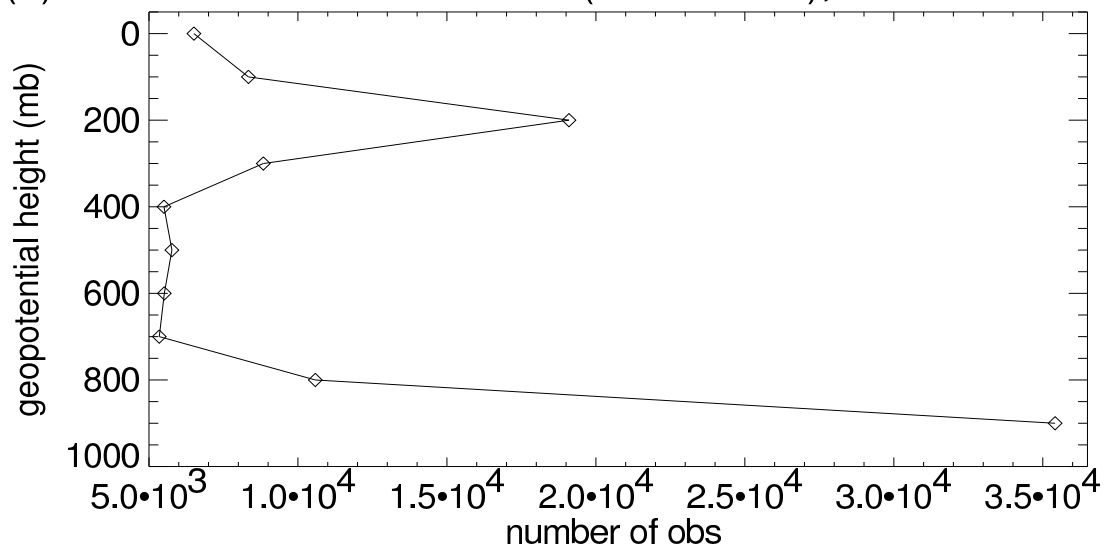## (b) Wind obs betwen diff levels(1900/01/03), total number:110917

Fig. 2. As in Fig.1, but for wind.

(a) Number of global observations

number of obs

$3.5 \cdot 10^5$
$3.0 \cdot 10^5$
$2.5 \cdot 10^5$
$2.0 \cdot 10^5$
$1.5 \cdot 10^5$
$1.0 \cdot 10^5$
$5.0 \cdot 10^4$

□ t00Z, ave = 284960
△ t06Z, ave = 199518
◇ t12Z, ave = 268501
✳ t18Z, ave = 241573

0   20   40   60   80   100   120

time(6-hour), from 011500 to 021518, 2003

(b) Number of S.H. observations

number of obs

$8 \cdot 10^4$
$6 \cdot 10^4$
$4 \cdot 10^4$
$2 \cdot 10^4$

□ t00Z, ave = 68790
△ t06Z, ave = 55202
◇ t12Z, ave = 66427
✳ t18Z, ave = 59030

0   20   40   60   80   100   120
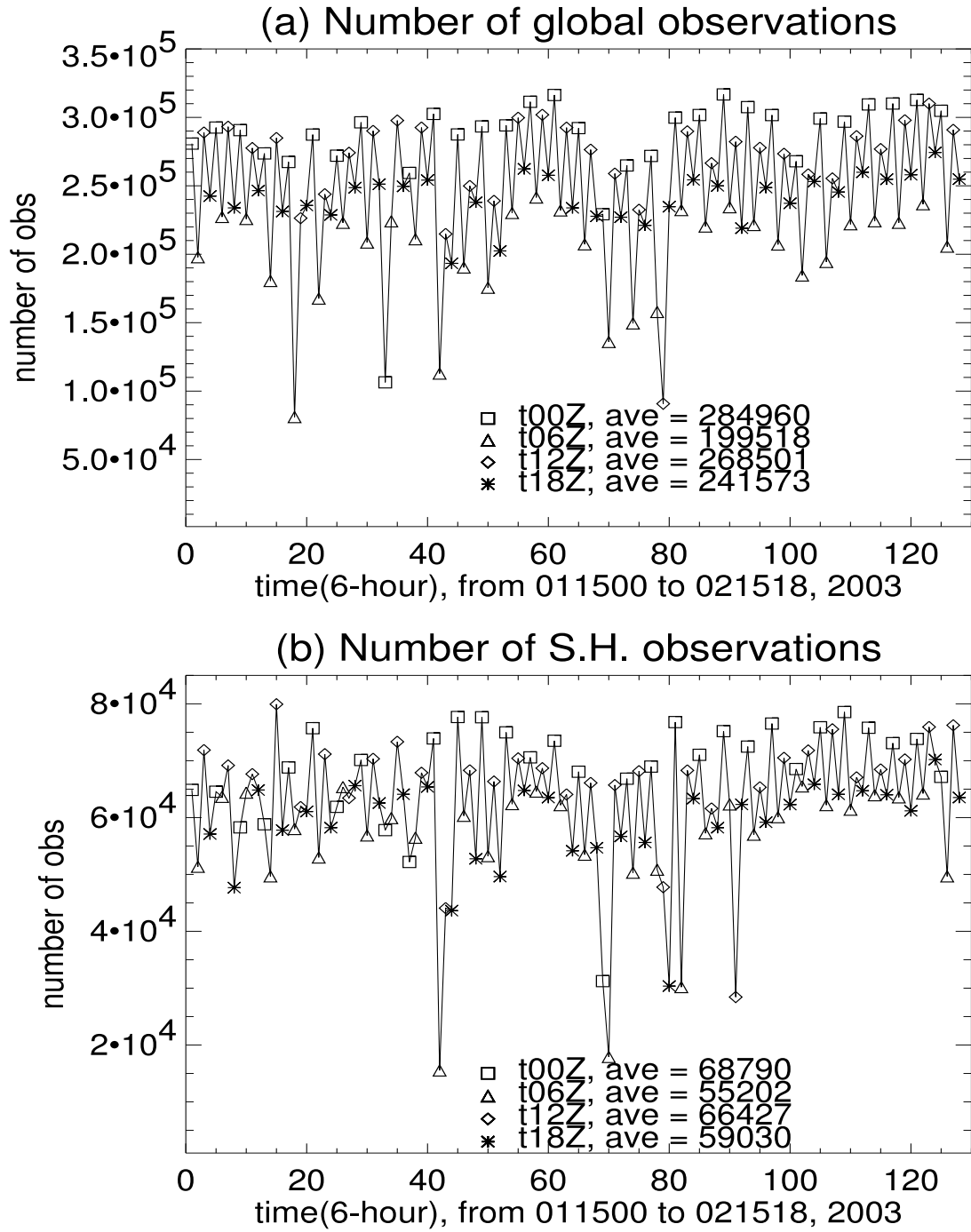
time(6-hour), from 011500 to 021518, 2003

Fig. 3. The number of observations at different cycles during the experimental period (a) over the globe; (b) over the Southern Hemisphere only.

(a) spread for energy(etkf)

(b) spread for energy(breeding)

(c) ratio of ana/fcst spread(etkf)
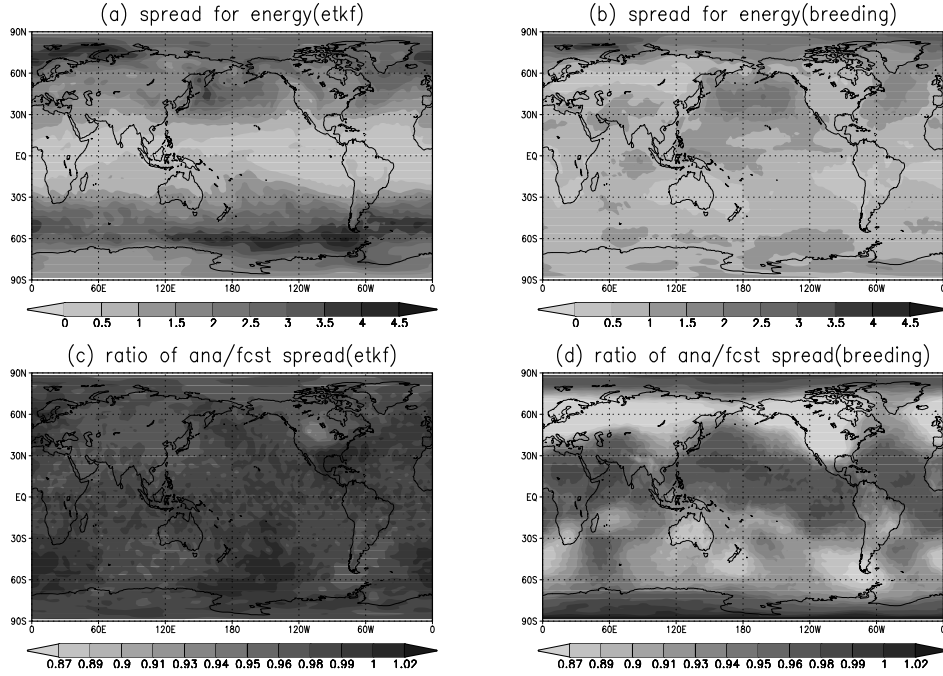
(d) ratio of ana/fcst spread(breeding)

Fig. 4. Vertically averaged global distribution of energy spread of analysis perturbations and the ratios of the analysis and forecast spread for both ETKF and breeding ensembles with (a) energy spread of ETKF; (b) energy spread of breeding; (c) ratio of analysis spread/forecast spread for ETKF; (d) ratio of analysis spread/forecast spread for breeding ensemble.

40

(a) ave over 0115-0215/03

ana energy perts
fcst energy perts
ratio of ana/fcst for energy

thick--> etkf
thin --> breeding

ave ana (etkf)=2.172
ave fcst (etkf)=2.222
ave ratio (etkf)=0.986

ave ana (breeding)=1.602
ave fcst (breeding)=1.694
ave ratio (breeding)=0.953

vertical pressure levels (mb)

Energy perts/ratio of ana/fcst

(b) ave over 0115-0215/03

ana energy perts
fcst energy perts
ratio of ana/fcst for energy

thick--> etkf
thin --> breeding

Latitude
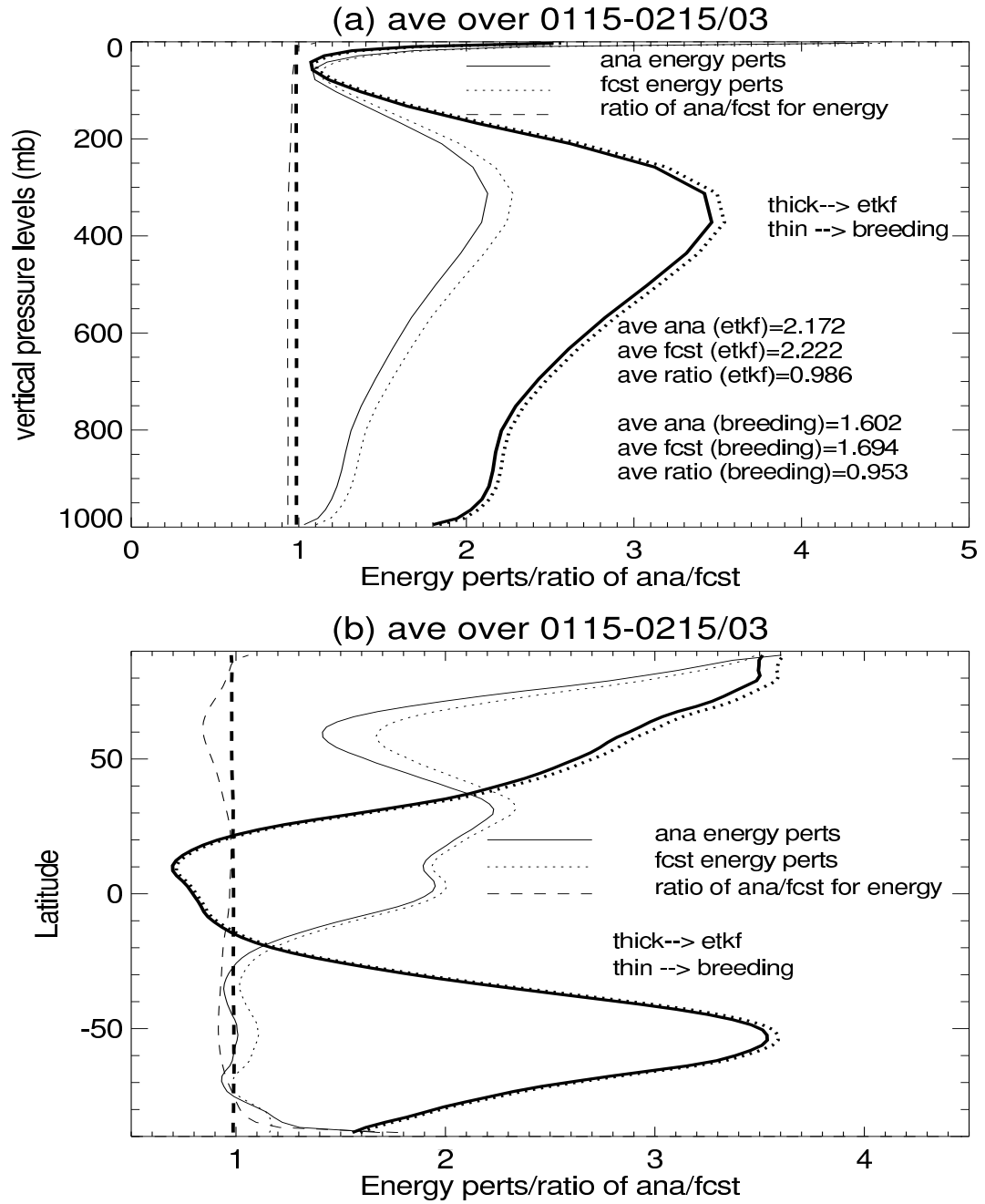
Energy perts/ratio of ana/fcst

Fig. 5. Energy spread distributions of ETKF (thick) and breeding (thin) ensemble perturbations (solid: analysis; dotted: forecast). The ratio of analysis/forecast perturbations is indicated by the dashed line. All the values are averaged over the period 01/15-02/15/2003, with (a) vertical distribution as a function of pressure; (b) distribution by latitude.
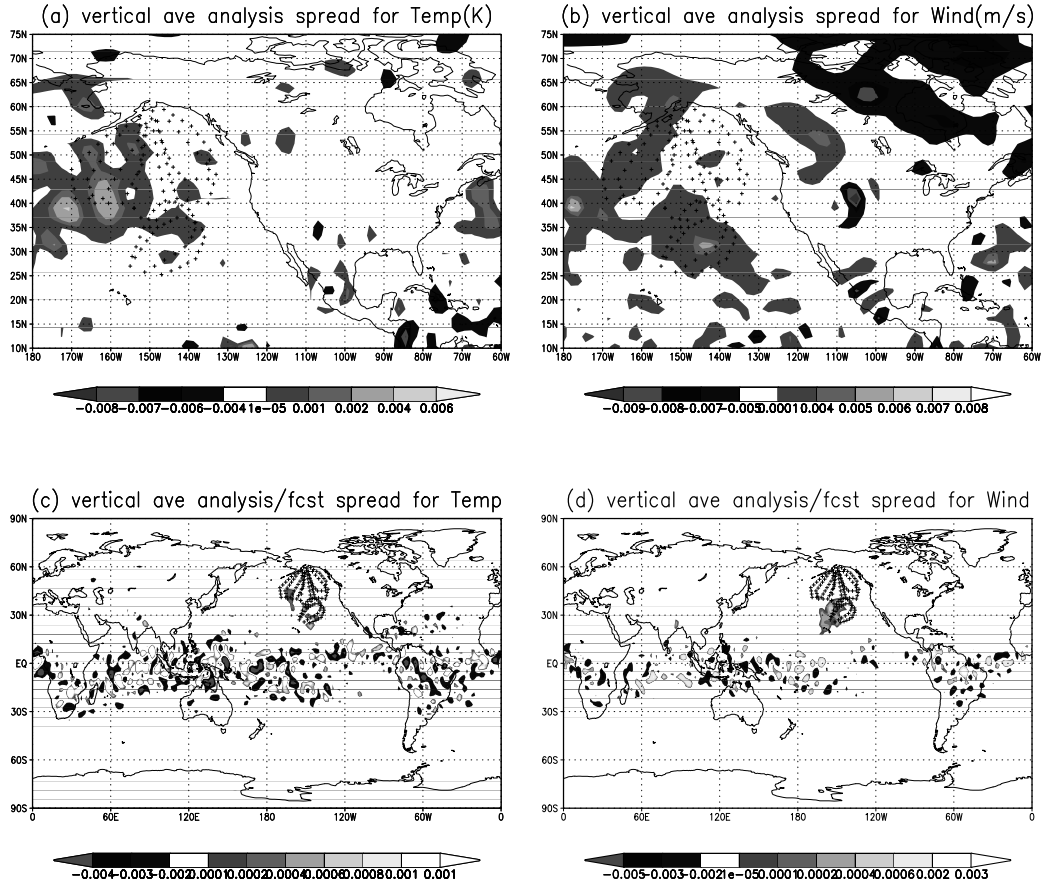
Fig. 6. The difference of vertically averaged analysis spread for (a) temperature and (b) wind, between two experiments with and without WSR data. The difference of analysis/forecast spread for (c) temperature and (d) wind, between the same two experiments.

(a) ave over 0115-0215/03 (ETKF)

dim of 10 fcst perts = 8.90

dim of the 1st 5 directions = 4.97

dim of 10 analysis perts = 9.00

(b) ave over 0115-0215/03(breeding)

dim of all 10 breeding fcst perts = 5.89
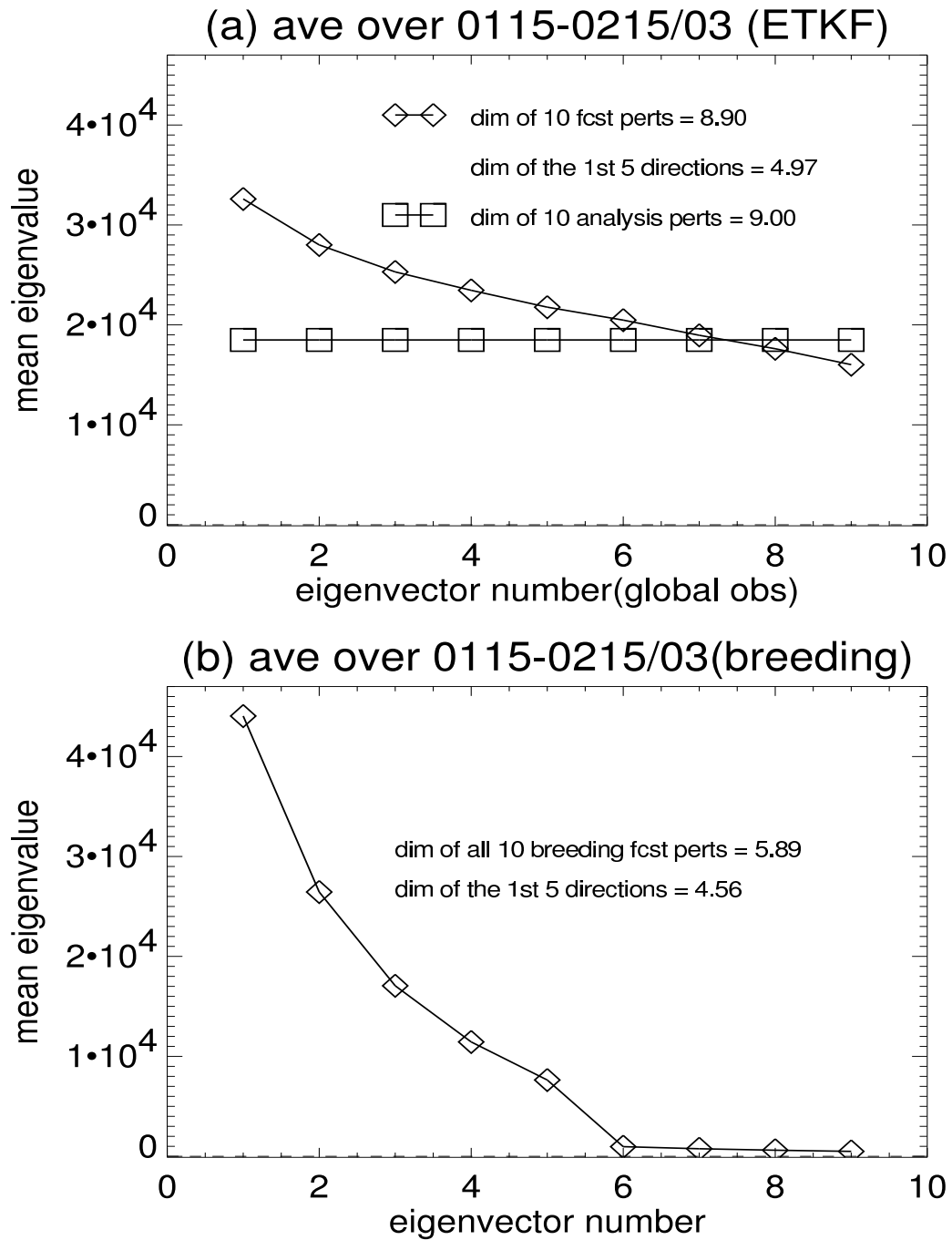
dim of the 1st 5 directions = 4.56

Fig. 7. The averaged variance distributions along different eigen-directions of forecast (diamond) and analysis (square) covariance matrices in the normalized observational space, for the: (a) ETKF ensemble; (b) breeding ensemble.

(a) Dim(0115-0215/03, 3-D Temp)

thick -> etkf
thin -> breeding

thick solid -> 5 ana perts, ave= 4.719
thick dotted-> 5 fcst perts, ave = 4.697

thin solid -> 5 ana perts, ave= 4.596
thin dotted -> 5 fcst perts, ave= 4.609

(b) Dim ave 0115-0215/03, (2-D Temp)

thick -> etkf
thin -> breeding

thick solid -> 5 ana perts, ave = 4.650

thick dotted -> 5 fcst perts, ave = 4.632

thin solid -> 5 ana perts, ave = 4.529
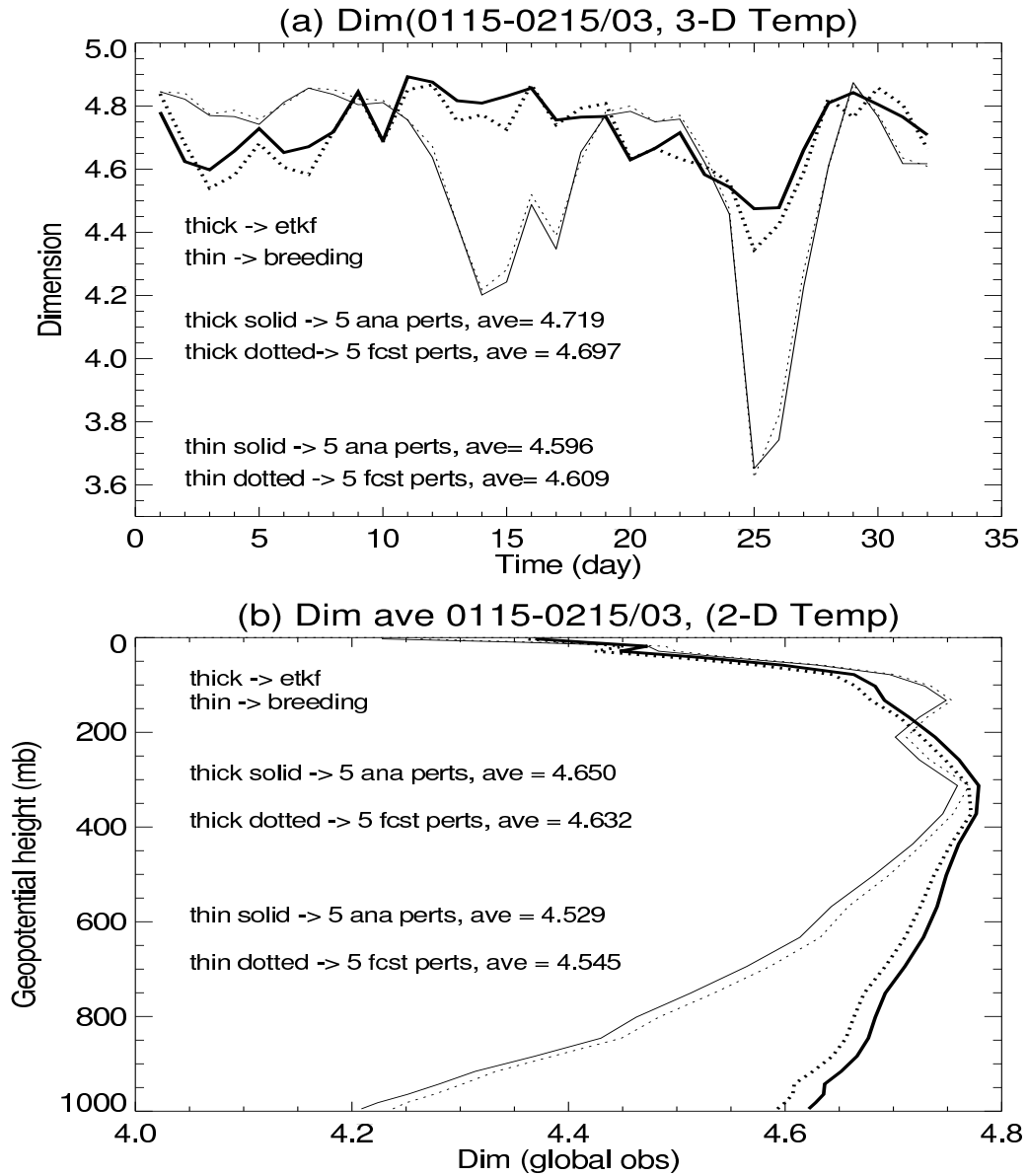
thin dotted -> 5 fcst perts, ave = 4.545

Fig. 8. The effective degrees of freedom of subspace spanned by 5 temperature perturbations (solid: 5 analysis perturbations; dotted: 5 forecast perturbations) from ETKF (thick) and breeding (thin) ensembles for: (a) EDF of subspace spanned by temperature perturbations in 3-dimensional grid point space, at different cycles during the period of experiments; (b) EDF of subspace spanned by 5 temperature perturbations in 2-dimensional grid point space at each pressure level.

(a) Ave local dim over global (5 temp perts)

square   -> ave dim of 5 ana perts for etkf
diamond -> ave dim of 5 ana perts for breeding

(b) Ave dim over global (5 temp perts)

thick->etkf
thin->breeding

solid: L=3
dotted: L=6
dashed: L=9
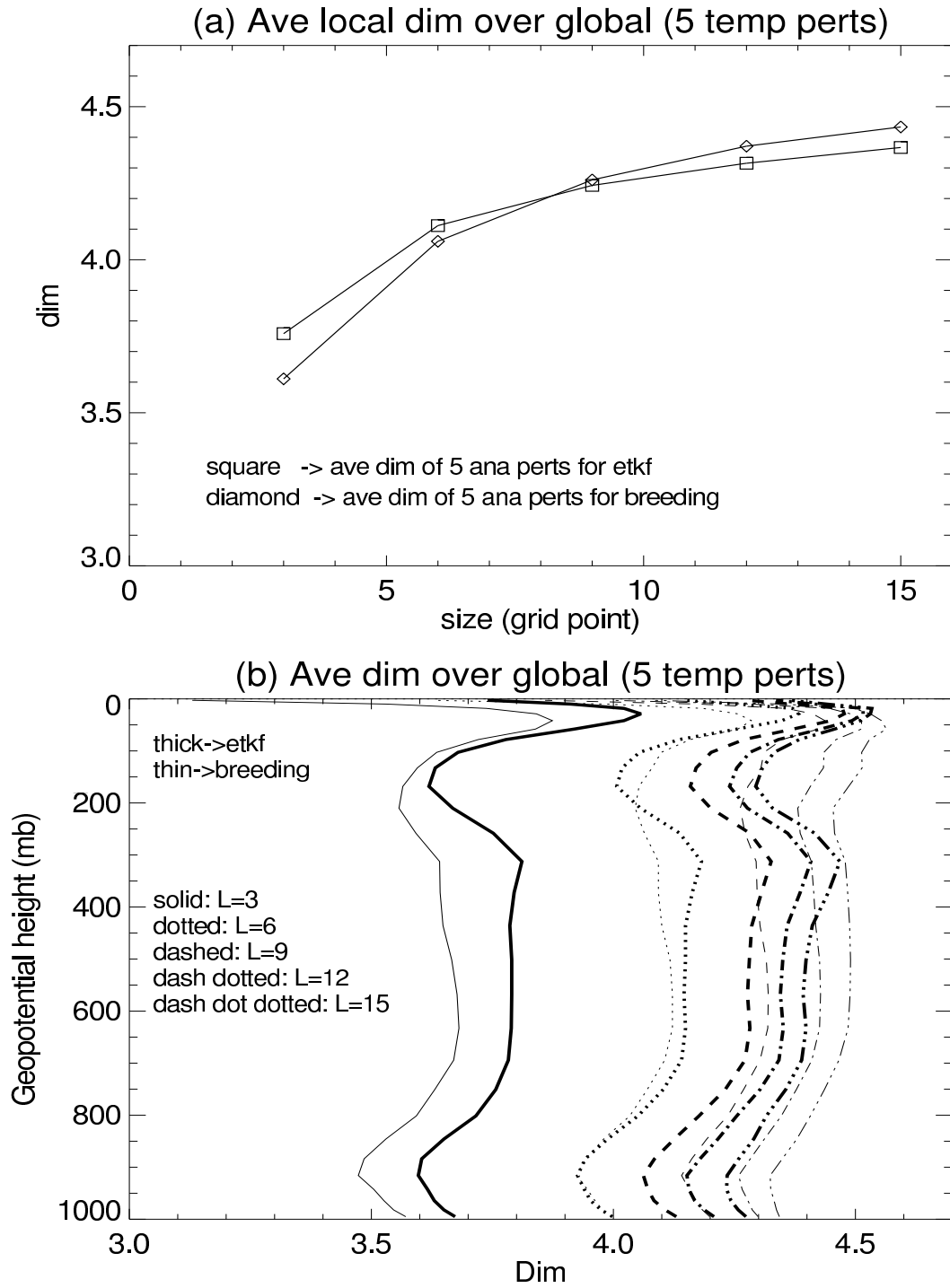dash dotted: L=12
dash dot dotted: L=15

Fig. 9. Local EDF for different number of grid points for: (a) averaged local
EDF as a function of L which is described in the text; and (b) local EDF at
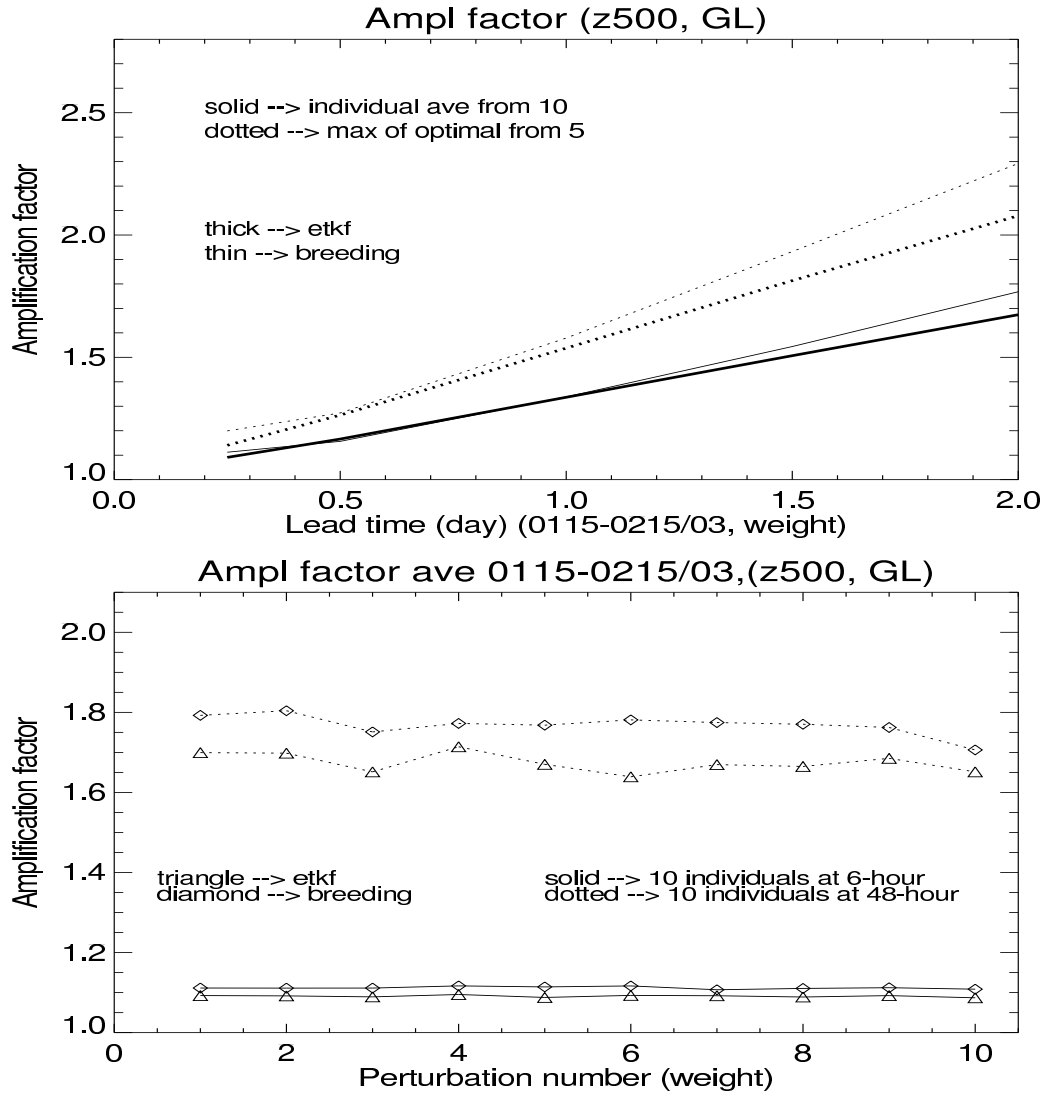each pressure level.

Ampl factor (z500, GL)

solid --> individual ave from 10
dotted --> max of optimal from 5

thick --> etkf
thin --> breeding

Amplification factor

Lead time (day) (0115-0215/03, weight)

Ampl factor ave 0115-0215/03,(z500, GL)

Amplification factor

triangle --> etkf
diamond --> breeding

solid --> 10 individuals at 6-hour
dotted --> 10 individuals at 48-hour

Perturbation number (weight)

Fig. 10. Amplification factors of ensemble perturbations for: (a) the averaged AF from 10 individual perturbations as a function of lead time (solid) and the maximum AF of optimally combined orthogonal perturbations from 5 original perturbations (dotted). (The results from ETKF and breeding ensembles are indicated by thick and thin lines respectively); and (b) the AF of 10 individual perturbations for forecast 2 lead times (solid: 6-hour, dotted: 48-hour), where the ETKF and breeding ensembles are indicated by triangles and diamonds, respectively.
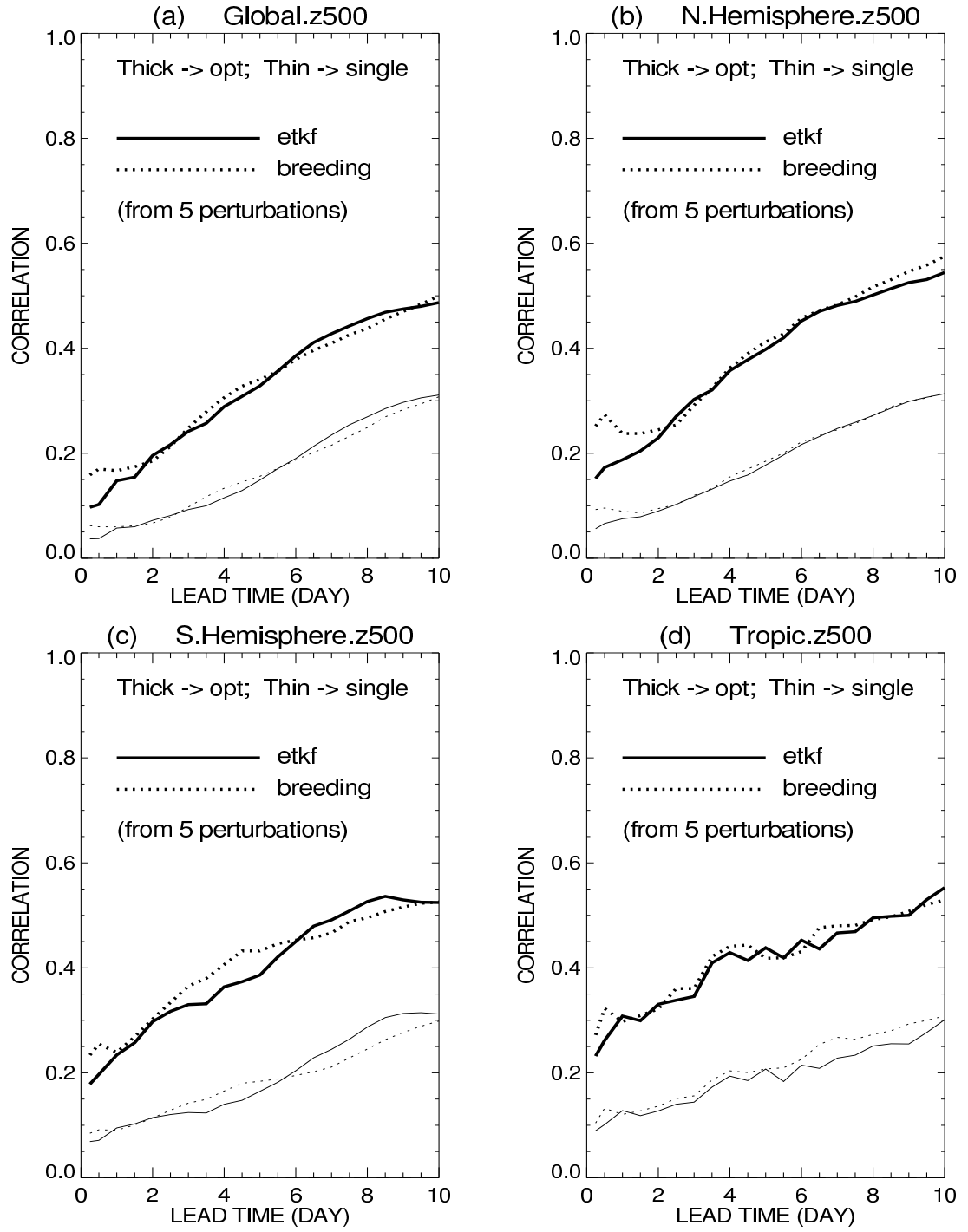
46

Fig. 11. The PECA values for ETKF (solid) and breeding (dotted) ensembles from 5 perturbations only. Shown in thick and thin lines are PECA from 5 optimally combined perturbations and average PECA from 5 individual perturbations.

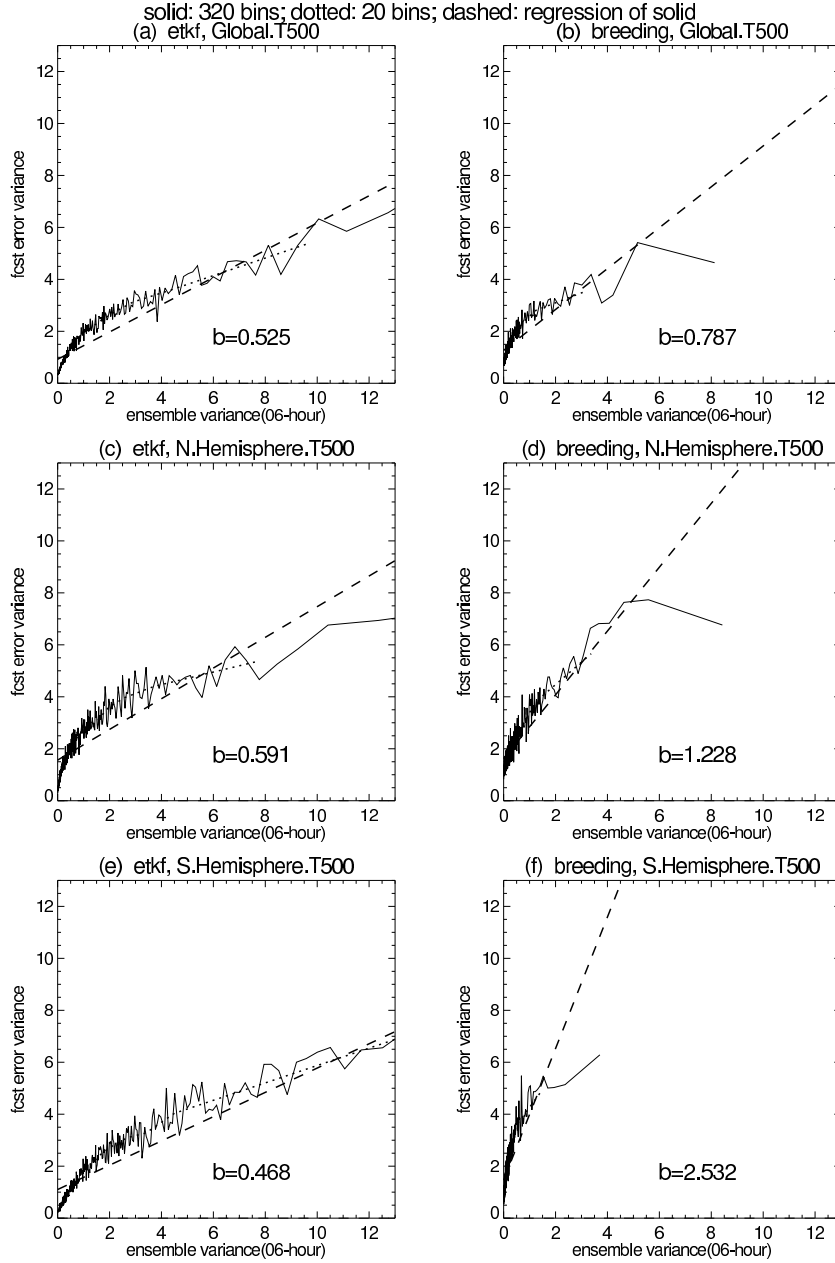solid: 320 bins; dotted: 20 bins; dashed: regression of solid

Fig. 12. Derived ensemble variance and forecast error variances at all grid points for 500mb temperature, for ETKF (left panel) and breeding (right panel); for global (top), Northern (middle) and Southern (bottom) Hemisphere regions. The average value from each of 320 bins is indicated by solid lines. Dotted lines show the results from 20 bins only. The linear regression line from 320 bins is displayed by a dashed line.
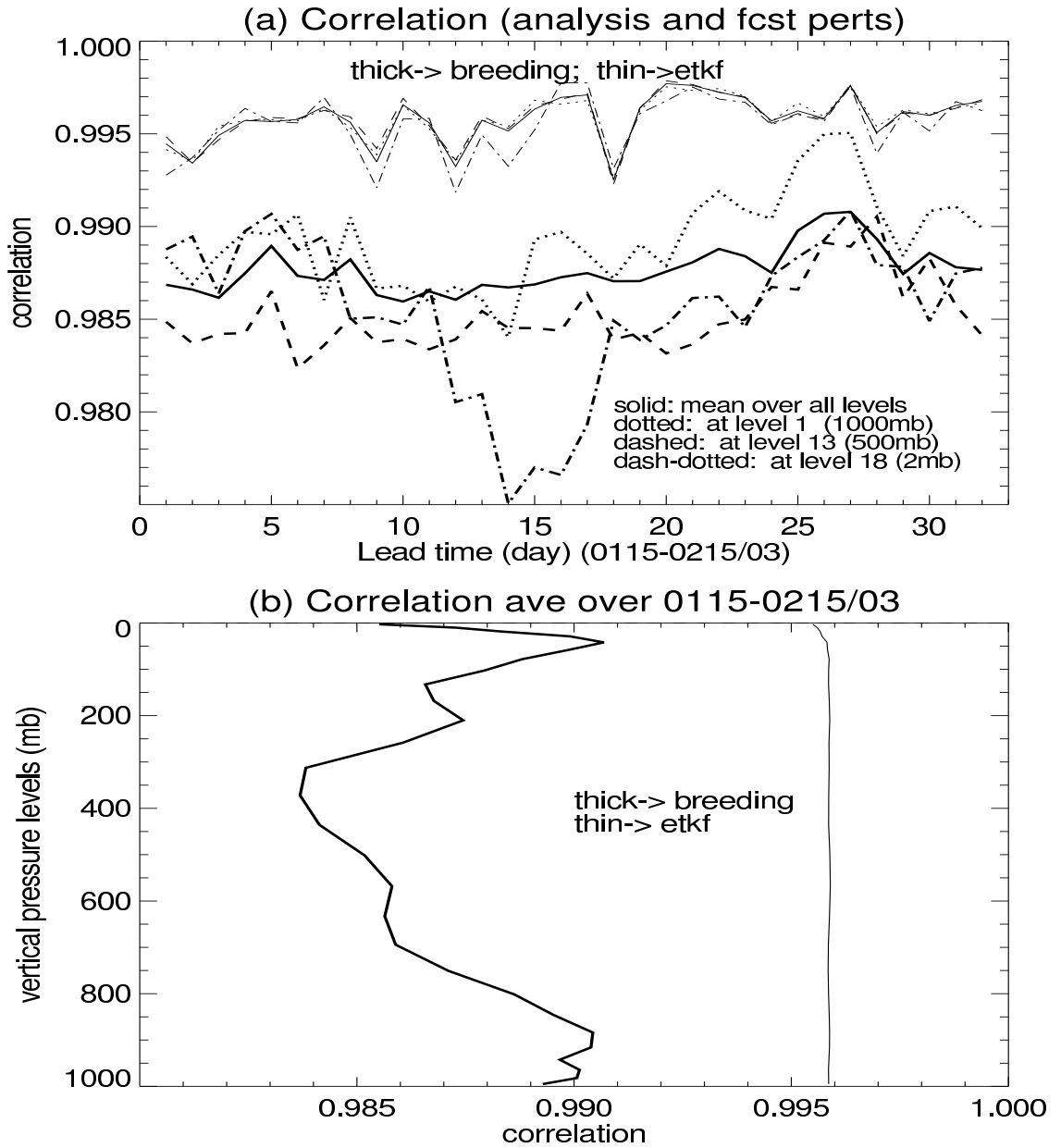
Fig. 13. Averaged correlation over 10 members between forecast and analysis perturbations for ETKF (thin) and breeding (thick) ensembles. (a) correlation as a function of time. The mean correlation over all levels, correlations at levels 1000mb, 500mb and 2mb are shown in solid, dotted, dashed and dash-dotted lines respectively. (b) vertical correlation distribution averaged over time.